

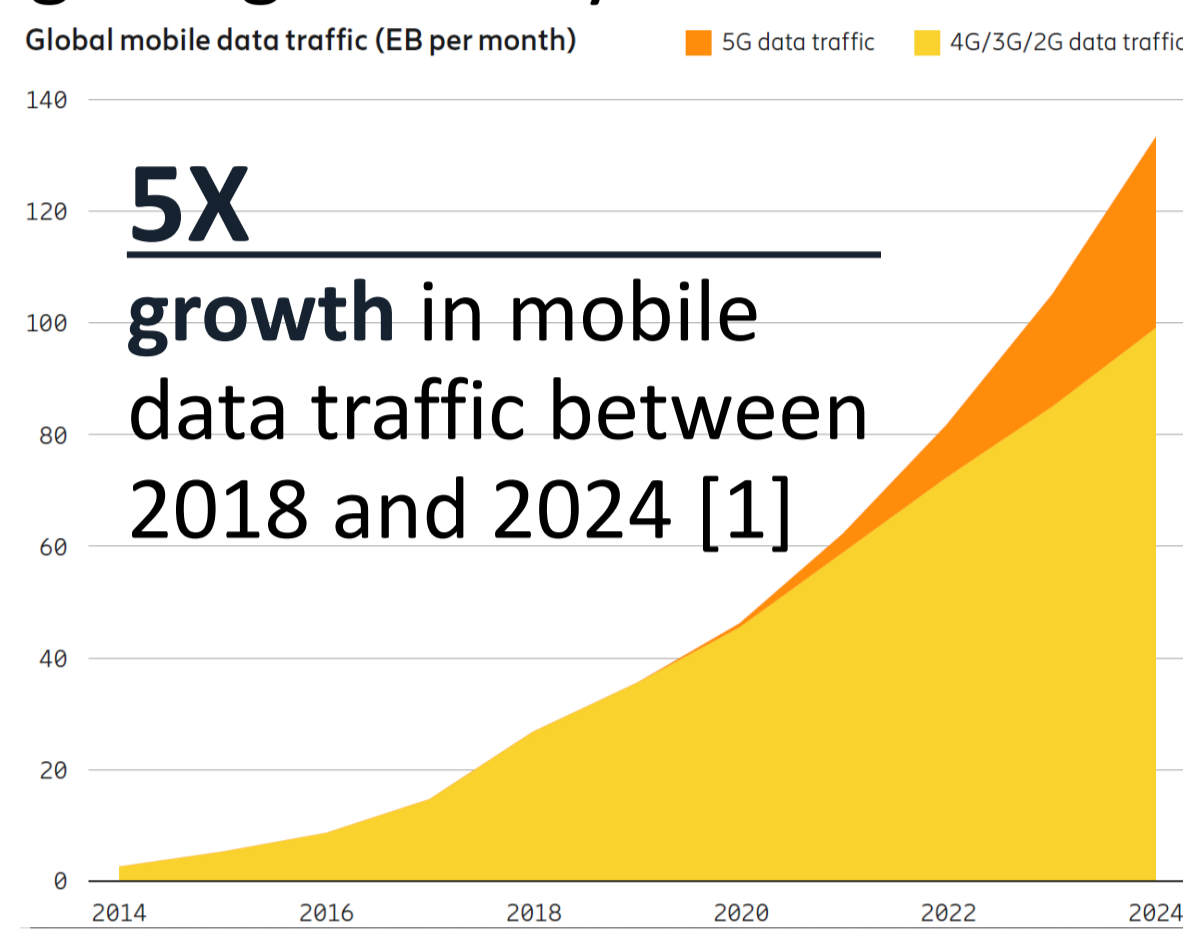
Antonio Montieri

Tutor: Antonio Pescapè

XXXII Cycle - II year presentation

Toward Effective Mobile Encrypted Traffic Classification through Deep Learning

Massive usage of mobile devices has changed significantly the network traffic



Traffic Classification (TC) of mobile apps

- provides valuable information for advertisement, insurance, safety, etc.
- raises privacy and security issues

Mobile Traffic Classification

Associating traffic classification objects to the mobile apps generating them



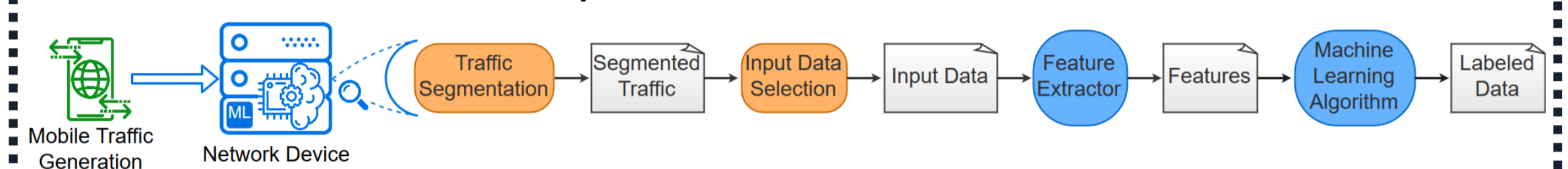
TC challenges are exacerbated

- Adoption of encrypted protocols
- Huge number of apps to discriminate
- Heterogeneous and dynamic nature of mobile traffic

Traditional Machine Learning (ML) classifiers

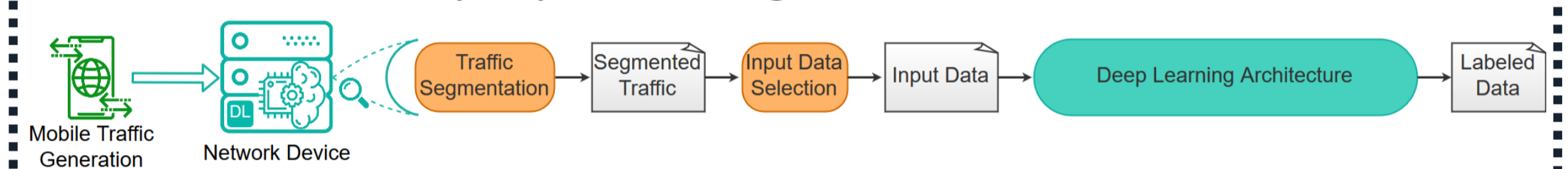
rely on domain-expert driven handcrafted features

- Time-consuming and unsuited to automation
- Outdated compared to the evolution of mobile traffic



Deep Learning (DL) classifiers are directly fed with input data

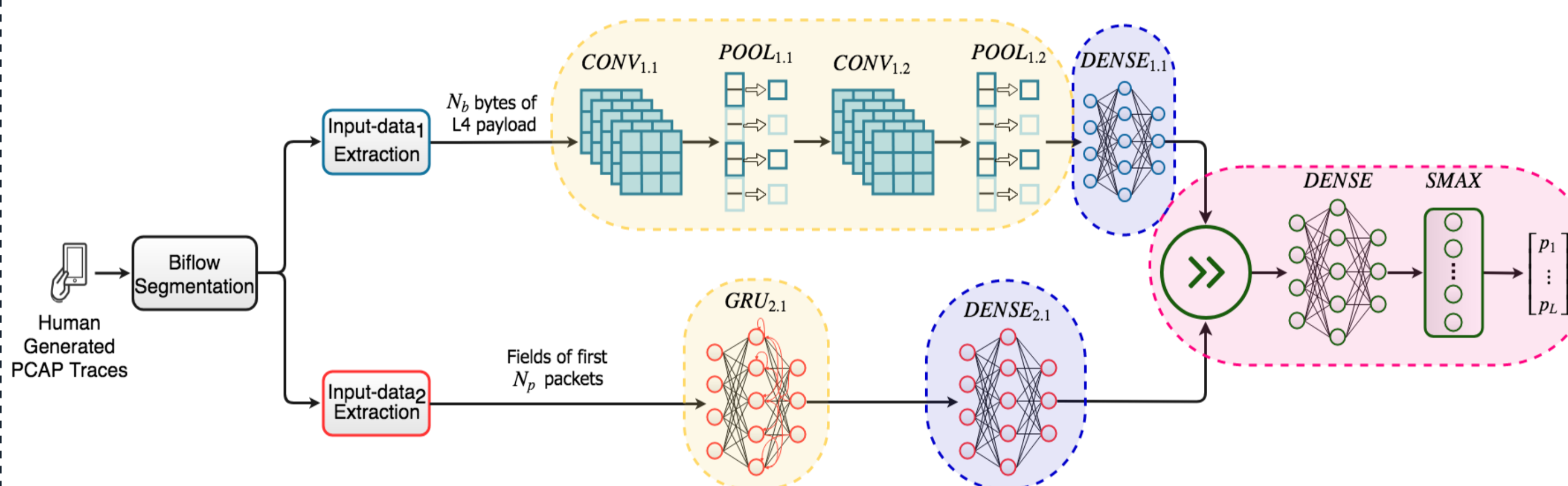
- Automatic hierarchical feature extraction
- Reduced preprocessing effort



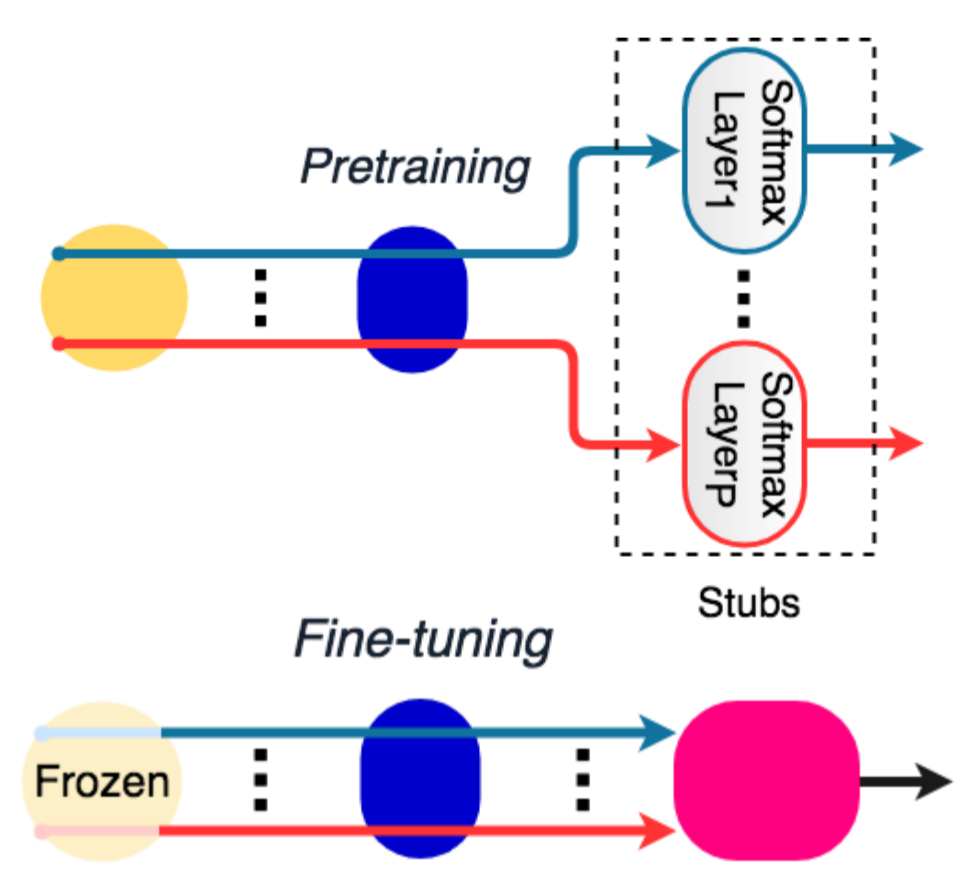
DL could be the stepping stone toward the achievement of high performance in mobile TC

Naïve adoption of DL implies misleading design choices and lead to biased conclusions [2] due to the peculiar nature of traffic data

Multimodal DL-based Mobile Traffic Classification (MIMETIC) framework has the capability of effectively exploiting the "multimodal" nature of the different views of a TC object [3]



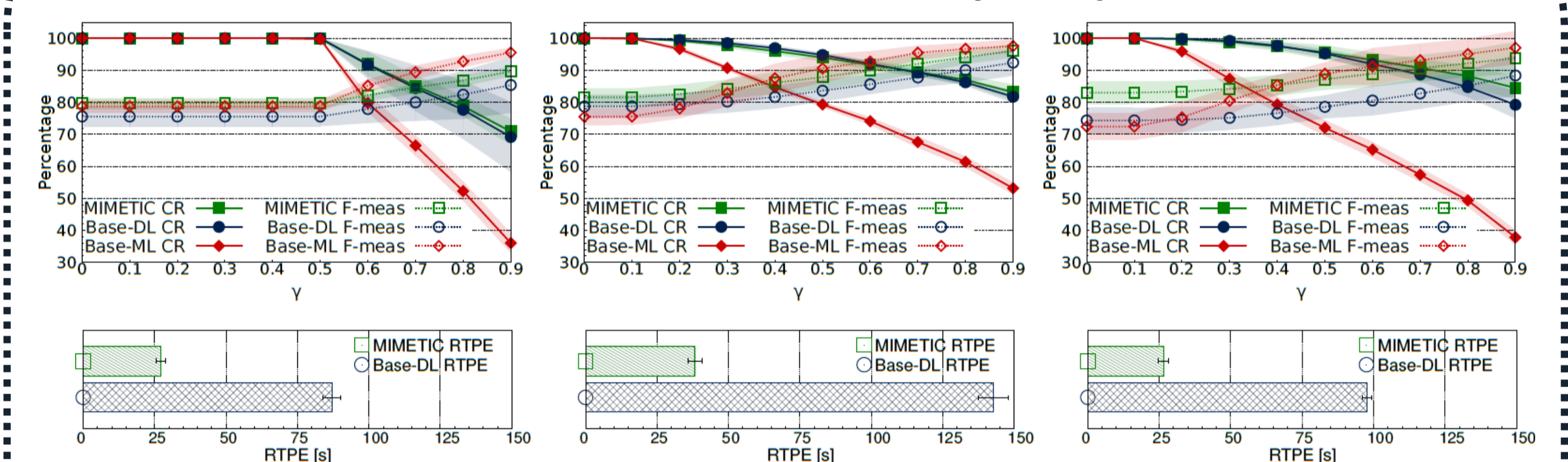
Proposed Training Procedure



Datasets

Platform	Dataset Description	Collection Period	Volume
FB/FBM	Binary, Collected @ ARCLAB DIE TI, > 100 users	05/17 - 03/18	31.0k biflows
Android	Multi-class, Collected by HUAWEI, 49 apps	04/15 - 01/17	55.5k biflows
iOS	Multi-class, Collected by HUAWEI, 45 apps	09/15 - 01/17	37.2k biflows

Mobile TC Performance vs. Reject Option



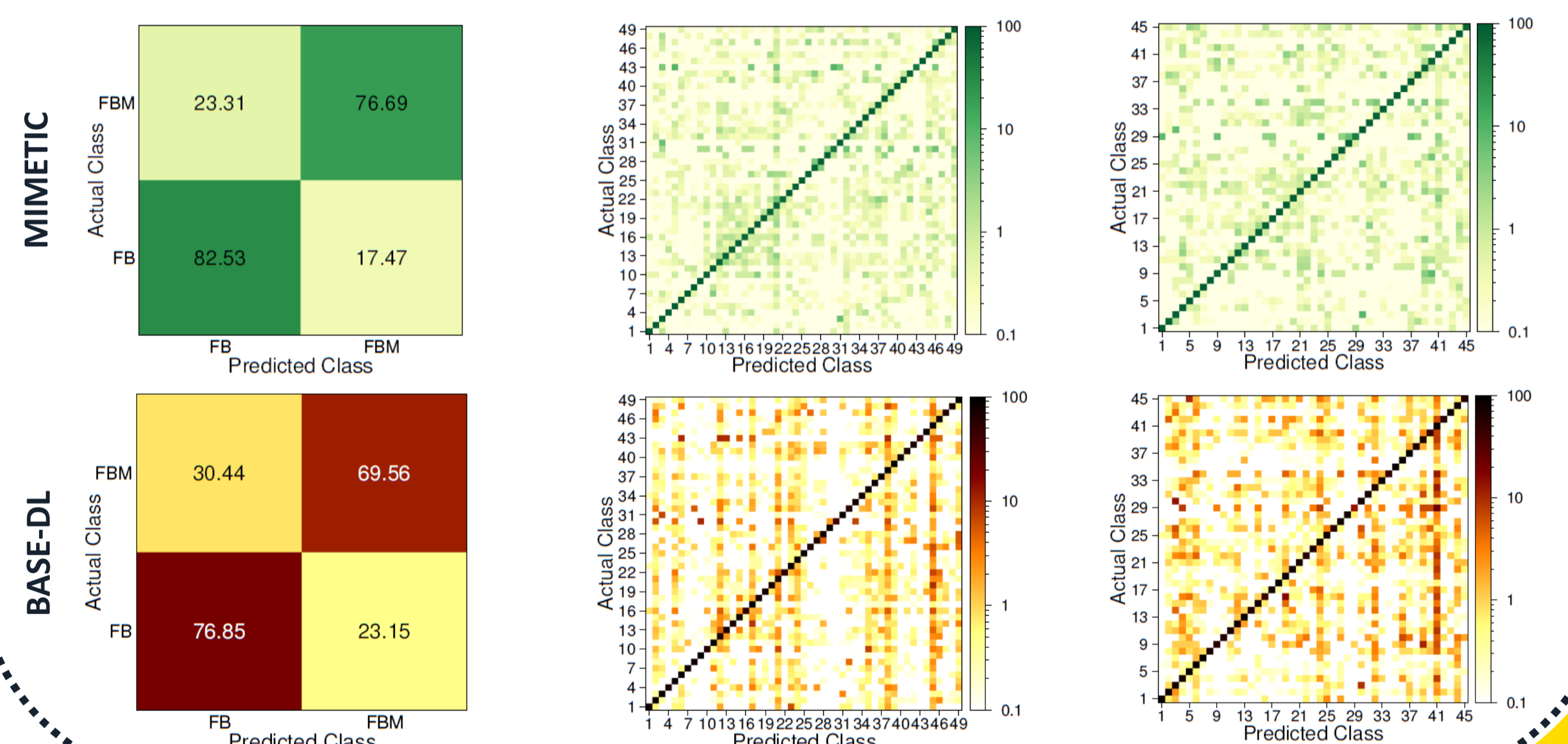
Comparison of MIMETIC framework with

- best single-modality DL classifier [2]
- state-of-the-art ML mobile-traffic classifier [4]

MIMETIC framework outperforms baselines

- Up to +8.58% F-measure improvement
- Run Time Per Epoch (RTPE) > 3.5x lower

Fine-grained Performance



Contacts

Antonio Montieri, Ph.D Student
 Phone: +39 081 76 83821
 Email: antonio.montieri@unina.it
 Skype ID: antmontieri



Traffic Research Group

<http://traffic.comics.unina.it>



Collaborations



[[simula](http://simula.research.laboratory). research laboratory]



Next Steps

- Evaluation of the purity of labeled samples used for training
- Exploitation of massive unsupervised data for improved learning
- Adoption of pre-trained architectures and more sophisticated DL layers
- Design of DL architectures able to cope with more challenging TC objects
- Adoption of Big Data paradigm to cope with increased training complexity

References

- [1] Fredrik Jejdling *et al.* "Ericsson mobility report", Ericsson AB, Business Area Networks, Stockholm, Sweden, Tech. Rep. EAB-18, 2018.
- [2] G. Aceto, D. Ciuonzo, A. Montieri, and A. Pescapè, "Mobile encrypted traffic classification using Deep Learning: Experimental evaluation, lessons learned, and challenges," IEEE Transactions on Network and Service Management, under revision, 2018.
- [3] G. Aceto, D. Ciuonzo, A. Montieri, and A. Pescapè, "MIMETIC: Mobile Encrypted Traffic Classification using Multimodal Deep Learning," IEEE Transactions on Mobile Computing, under revision, 2018.
- [4] V. F. Taylor, . Spolaor, M. Conti, and I. Martinovic, "Robust smartphone app identification via encrypted network traffic analysis," IEEE Transactions on Information Forensics and Security, vol. 13, no. 1, pp. 63–78, 2018.