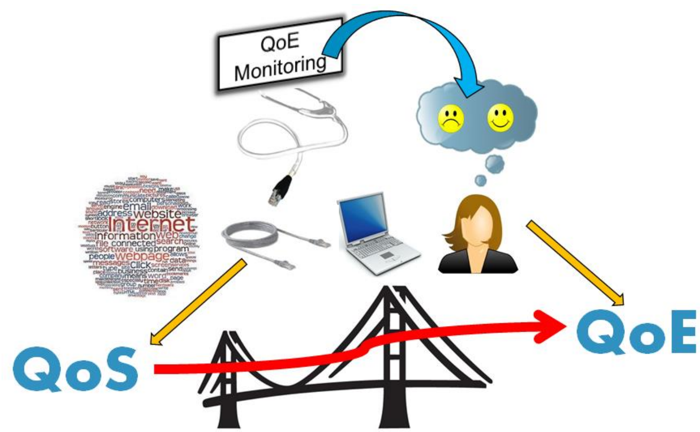
Improving QoE Prediction in Mobile Video through Machine Learning

P. Casas, S. Wassermann

AIT Austrian Institute of Technology, Vienna

University of Würzburg, Institute of Computer Science, Würzburg



AGENDA

- ***Brief State of the Art in Mobile Video Modeling***
- *Data Description*
- *Machine Learning Models for QoE Prediction*
- *QoE Modeling Results*
- *Concluding Remarks*

QoE in Video Streaming for Smartphones

- **Video Streaming QoE** is mainly **affected by stalling** (i.e., re-buffering events) and **video quality switches** (HAS – HTTP Adaptive Streaming)

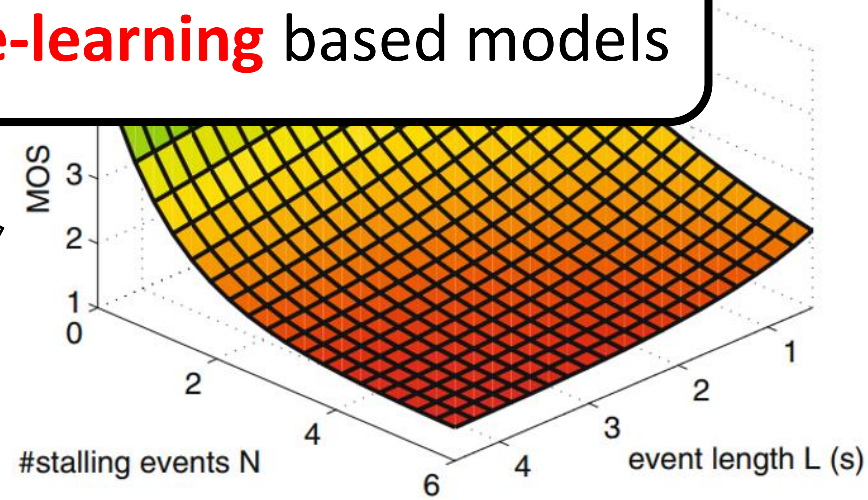
- **Initial playback delay** has a **limited impact on QoE**

- **In smartphones**, where displays are rather small w.r.t standard devices, **video quality switches do not have an important impact** on the perception of the user [*]

- **Most**
  -
  - ***L***: average stalling duration

We propose a **new model for QoE in Video Streaming for Smartphones**, using **machine-learning** based models

$$\boxed{MOS = a.e^{-(b.L + c).N}}$$



[*] Exploring QoE in Cellular Networks: How Much Bandwidth do you Need for Popular Smartphone Apps?
P. Casas et al., ACM SIGCOMM All Things Cellular Workshop 2015

AGENDA

- *Brief State of the Art in Mobile Video Modeling*
- ***Data Description***
- *Machine Learning Models for QoE Prediction*
- *QoE Modeling Results*
- *Concluding Remarks*

Dataset Overview (1/3)

The University of Texas at Austin
LIVE
Laboratory for Image & Video Engineering

- Train and test different **regression models** mapping video stalling patterns into QoE (MOS)

- Publicly available **subjective QoE measurements dataset** → ***LIVE-Avvasi Mobile Video database*** (University of Texas @Austin):

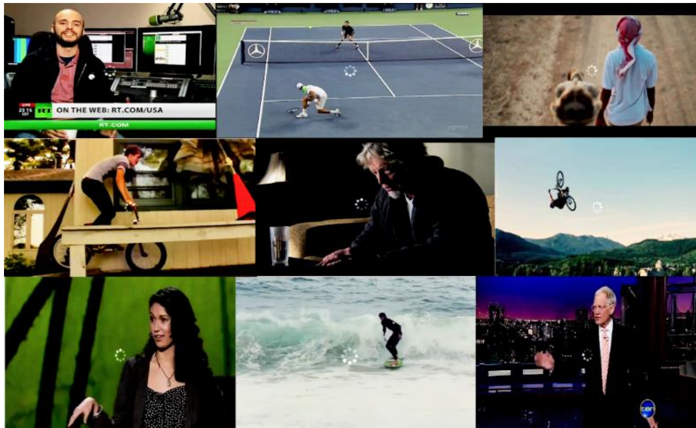
  - **174 distorted videos** generated from 24 reference videos with **26 unique stalling events**

  - **4830 ratings obtained from 54 subjects** who viewed the videos in smartphones

  - Ratings correspond to **MOS scores in ACR scale** (1 – *bad quality* to 5 – *excellent quality*)

  - reference videos: **HD YouTube and Vimeo**, with a **duration range 30s to 2min**

  - **different contents:** sports, documentaries, advertisement, music clips



Dataset Overview (2/3)

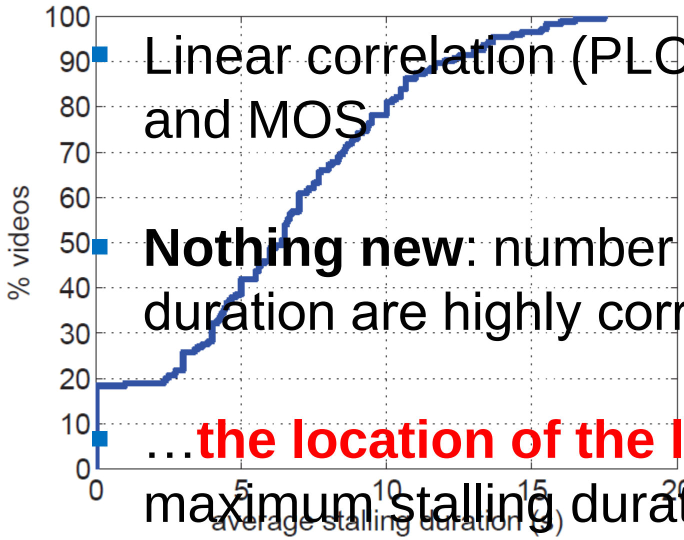
- We extract **19 different features characterizing the stalling patterns** undergone by the videos, including:
  - number and frequency of stalling events
  - initial playback delay
  - duration of stallings
  - **location of stallings within the video stream**
  - particular video contents (e.g., frames per second)

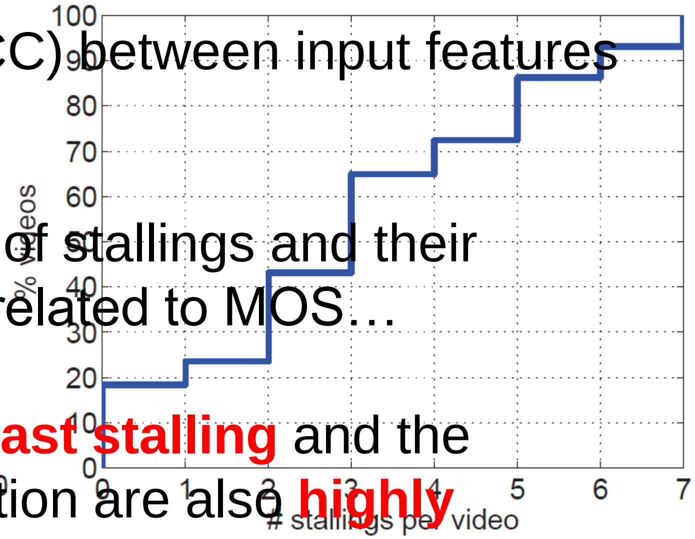
Highly relevant, due to memory effects!!!

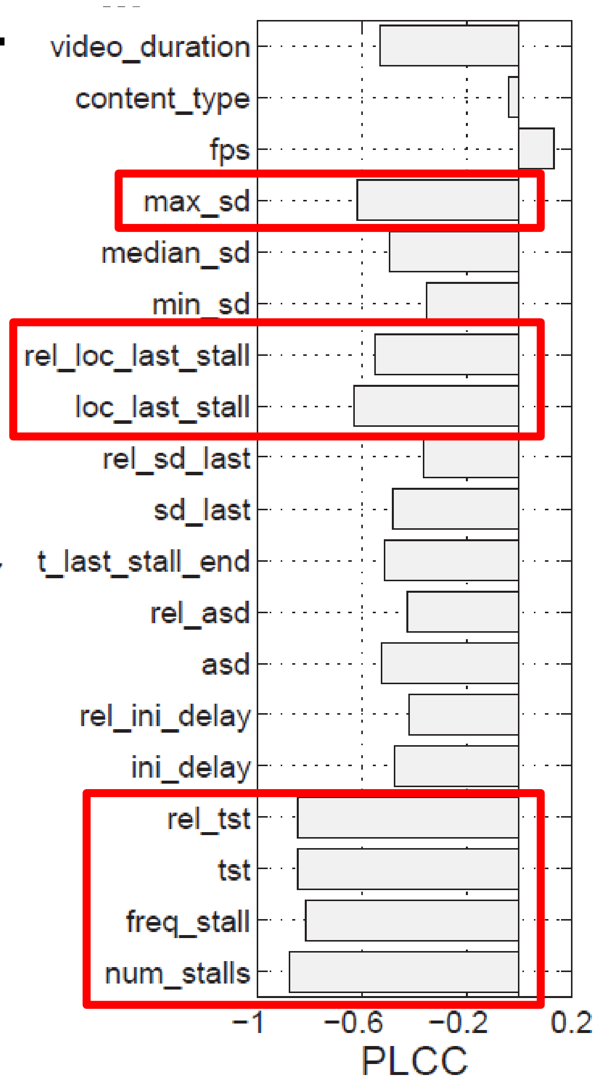
	feature	description
f_1	num_stalls	total number of stallings
f_2	freq_stalls	frequency of stallings
f_3	tst	total stalling time
f_4	rel_tst	tst, relative to video duration
f_5	ini_delay	initial playback delay
f_6	rel_ini_delay	ini_delay, relative to video duration
f_7	asd	average stalling duration
f_8	rel_asd	asd, relative to video duration
f_9	t_last_stall_end	elapsed time between end of last stalling and end of the video
f_{10}	sd_last	duration of last stalling
f_{11}	rel_sd_last	sd_last, relative to video duration
f_{12}	loc_last_stall	elapsed time between start of the video and start of last stalling
f_{13}	rel_loc_last_stall	loc_last_stall, relative to video duration
f_{14}	min_sd	minimum stalling duration
f_{15}	median_sd	50%-percentile of stalling duration
f_{16}	max_sd	maximum stalling duration
f_{17}	fps	video frames per second
f_{18}	content_type	video category (e.g., sports, news, etc.)
f_{19}	video_duration	total length of the video
	MOS	average video MOS score

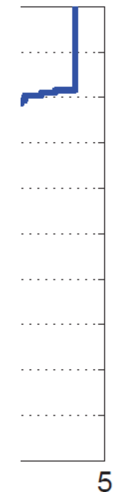
Dataset Overview (3/3)

- **Temporal features** → both **absolute and relative** (to video length) **values**
- Empirical distributions: **avg. stalling duration, num.**
- Linear correlation (PLCC) between input features and MOS
- **Nothing new**: number of stallings and their duration are highly correlated to MOS…
- …**the location of the last stalling** and the maximum stalling duration are also **highly correlated to MOS** (**new!**)
- Can we **exploit these new features** to improve QoE prediction results? **YES!!!**







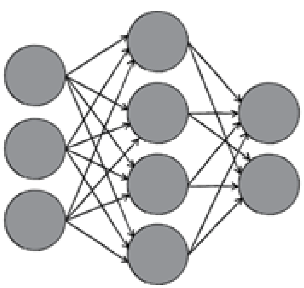


AGENDA

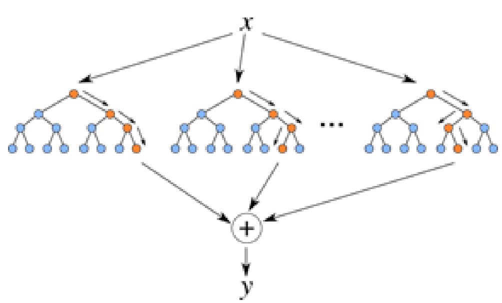
- *Brief State of the Art in Mobile Video Modeling*
- *Data Description*
- ***Machine Learning Models for QoE Prediction***
- *QoE Modeling Results*
- *Concluding Remarks*

Machine Learning Models

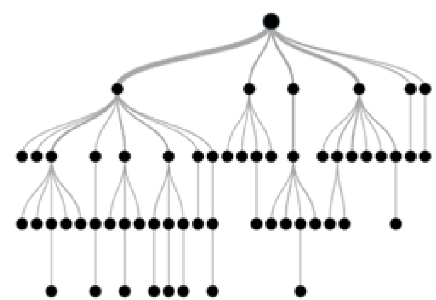
- **Supervised ML models (regression)** to predict MOS for each video
- Different algorithms **trained on subjective test dataset** (10-fold cross validation)
- Benchmark **11 different learning models**:
  - Support Vector Machines (SVM)
  - Multiple classes of Decision Trees: random tree, Random Forest (RF), bagging, **continuous tree (M5P)**, Decision Stump (DS), discrete tree
  - Neural Networks (MLP)
  - Locally Weighted Learning (LWL)
  - Linear and Pace Regression
- **WEKA used as ML library**, grid search for parameter configuration



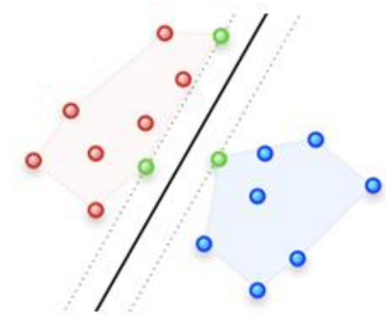
MLP model



RF model



C4.5 model



SVM model

AGENDA

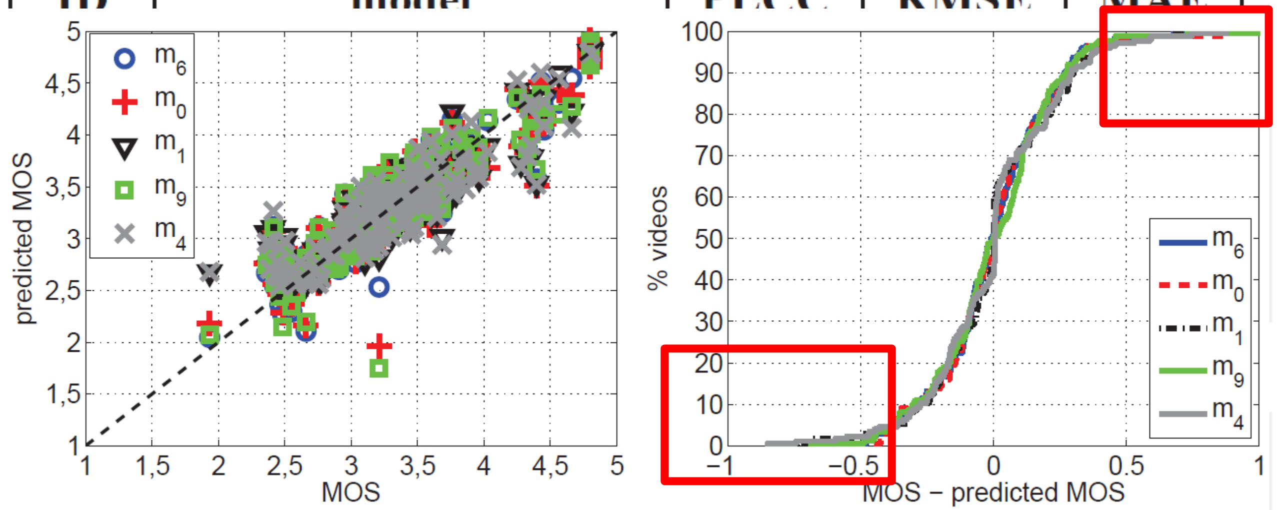
- *Brief State of the Art in Mobile Video Modeling*
- *Data Description*
- *Machine Learning Models for QoE Prediction*
- ***QoE Modeling Results***
- *Concluding Remarks*

Machine Learning Models Benchmarking

- **ML models benchmarking** using 3 performance metrics: linear correlation between predicted and real MOS (**PLCC**), root mean squared error (**RMSE**) and mean absolute error (**MAE**)

- **Top 5 models** include **SVM, RF, bagging tree, pace regression,** and **M5P**

- **Very high correlation** between predicted and real MOS (~0.95) and **limited prediction errors** (below 0.2 in a 5-points MOS scale)



(a) $\widehat{MOS}$ vs MOS. (b) distribution of MOS - $\widehat{MOS}$.

ID	model	PLCC	RMSE	MAE
m_7	linear regression	0.878	0.362	0.310
m_8	additive regression DS	0.921	0.301	0.231
m_9	**pace regression**	**0.948**	**0.242**	**0.179**
m_{10}	discrete regression tree	0.911	0.315	0.224

- A **negligible fraction of videos with prediction errors above 0.5**

- **M5P model selected as reference model for QoE modeling in mobile video** (due to its **simplicity** and **input/output visibility** – see next slide)

Overview of M5P Model

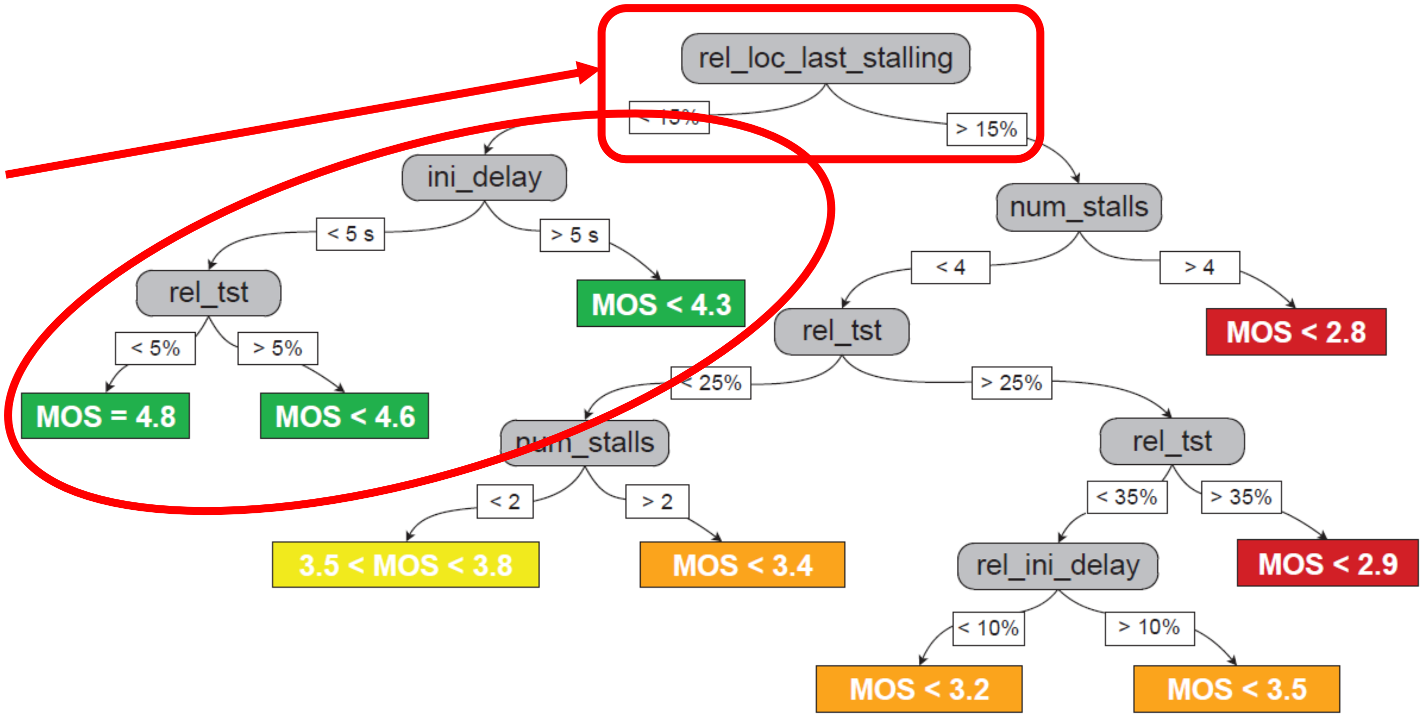
- The trained **M5P model** selects **5 out of the 19 input features**, including:
  - number of stallings – **num_stalls**
  - total stalling time (relative to video duration) – **rel_tst**
  - initial playback delay (absolute and relative) – **ini_delay** and **rel_ini_delay**
  - **location of last stalling event (relative to video duration) – rel_loc_last_stalling**

State of the Art

NEW!

- Approximated M5P decision tree (based on discretization)

- The **location of the last stalling event is critical** (tree root)
- If all **stallings occur in the first 15% of the video playback**, their **impact is almost negligible** (comparable to initial playback delay)

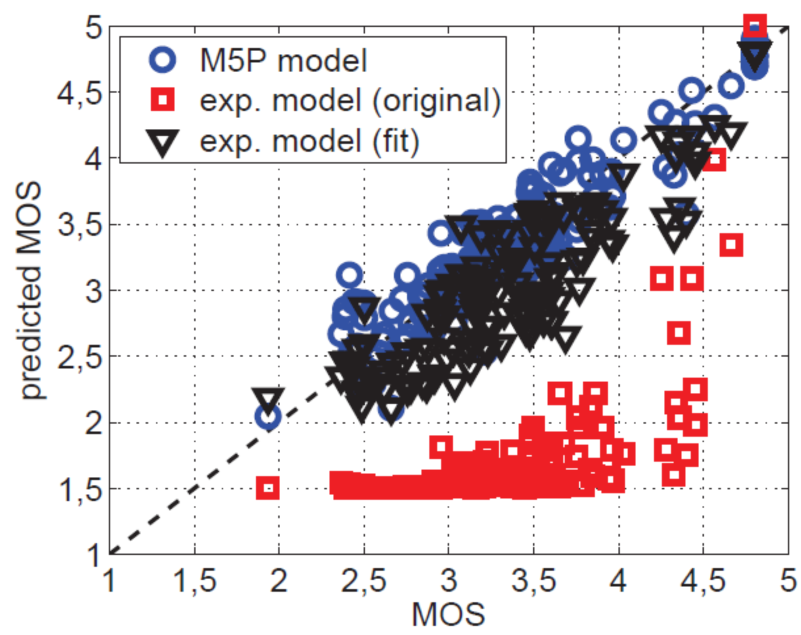


M5P vs. State of the Art Models

- We **compare our M5P model** with three **state-of-the-art** models for video streaming QoE prediction:
  - **exponential model**, using original parameters (***exp. original***) and those fitted to evaluation dataset (***exp. fit***)
  - non-linear, **filter-based model** with memory (***HW****) → model from group generating current dataset under study
  - **state machine-based model** (***DQS*****)

- **M5P clearly outperforms state-of-the-art models**
- much higher correlation and smaller errors
- exponential model @2nd place, after a careful re-calibration of its underlying parameters

m	E
M	57
exp. (or	51
ex	23
HV	
DQ	



$\widehat{\text{MOS}}$ vs MOS.

[*] D. Ghadiyaram et al., "**A Time-varying Subjective Quality Model for Mobile Streaming Videos with Stalling Events**", in *SPIE Optical Engineering Applications*, 2015.
[**] H. Yeganeh et al., "**Delivery Quality Score Model for Internet Video**", in *IEEE International Conf. on Image Proc.*, 2014.

AGENDA

- *Brief State of the Art in Mobile Video Modeling*
- *Data Description*
- *Machine Learning Models for QoE Prediction*
- *QoE Modeling Results*
- ***Concluding Remarks***

Conclusions

- We have **introduced a novel machine learning based model** for **multi-dimensional QoE prediction in mobile video streaming**.
- Based on decision trees, the **proposed model outperforms** previously proposed **state-of-the-art** models by **reducing prediction errors between 25% and almost 50%**.
- The proposed M5P model shows that **there is a clear influence of other stalling pattern descriptors generally neglected in previous work**…
- …in particular those linked to **the occurrence of the last stalling event**.
- The **M5P model** could **enhance current measurement tools and systems** for video streaming QoE prediction, suggesting **novel metrics** to measure in the future.
- We're currently **working on the generalization of the presented results**, considering **other datasets**

Thanks for Your Attention!