IEEE INTERNET OF THINGS JOURNAL, VOL. 6, NO. 1, FEBRUARY 2019 997

# Chaac: Real-Time and Fine-Grained Rain Detection and Measurement Using Smartphones

Hansong Guo, He Huang, *Member, IEEE*, Yu-E Sun, Youlin Zhang,
Shigang Chen, *Fellow, IEEE*, and Liusheng Huang

H. Guo is with the Department of Computer Science and Technology, University of Science and Technology of China, Hefei 230027, China, and also with the Department of Computer and Information Science and Engineering, University of Florida, Gainesville, FL 32611 USA (e-mail: guohanso@mail.ustc.edu.cn).

H. Huang is with the School of Computer Science and Technology, Soochow University, Suzhou 215006, China (e-mail: huangh@suda.edu.cn).

Y.-E Sun is with the School of Rail Transportation, Soochow University, Suzhou 215006, China (e-mail: sunye12@suda.edu.cn).

Y. Zhang and S. Chen are with the Department of Computer and Information Science and Engineering, University of Florida, Gainesville, FL 32611 USA (e-mail: ylzh10@ufl.edu; sgchen@cise.ufl.edu).

L. Huang is with the Department of Computer Science and Technology, University of Science and Technology of China, Hefei 230027, China (e-mail: lshuang@ustc.edu.cn).

# Outline

- Background and Motivations
- Previous Works
- System Description
- Evaluation
- Discussions and Conclusions

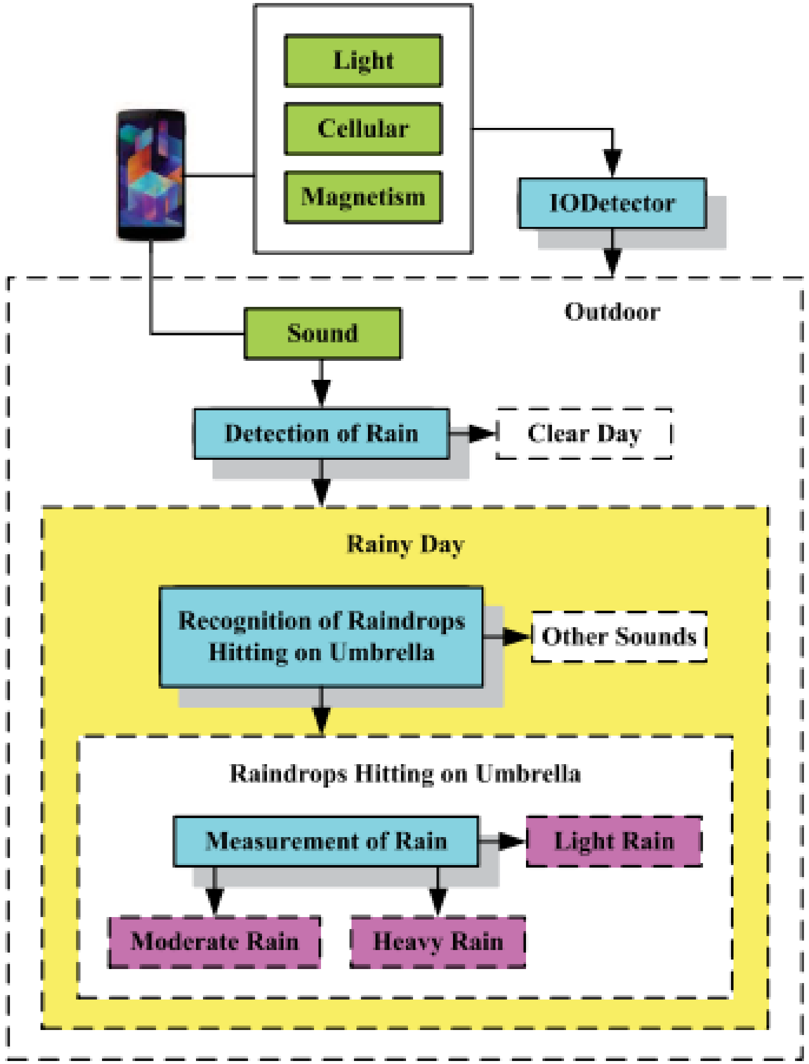
# Background and Motivations

- Existing rain gauge covers <1% of earth surface
- Existing weather apps have coarse granularity of temporal and spatial resolution
    - limited rain gauges that record rain data periodically
    - intensity of rain changes dramatically with time and space
- Chaac: a novel mobile sensing system that exploits opportunistically crowdsourced audio clips from smartphones
    - claimed to be the first work to detect and measure the intensity of rain by smartphone
    - real-time
    - fine-grained
- Evaluation results shows 92% and 93.9% accuracy for rain detection and rain measurements
    - claimed to be superior than any other relevant existing approach in both industrial and academic communities

# Previous Works of Rain Detection and Measurement

- Images-based
    - smartphone-based or in-vehicle camera –based
    - requires extra work from participants
    - in-vehicle camera lack the capacity of processing the images, need to transmit to server, rain detection are not real-time
- Wireless Links-based
    - cellular links of 38-GHz
    - not fine grained detection
- Satellite-based
    - combine rain radar and satellites by training a ANN
    - expensive, consumes a lot of energy, not fine-grained
- Wiper-based
    - crowdsourcing speed of vehicle wipers
    - wipers as rain detection sensors
    - only detects rain on the road, not fine grained
- Chaac:
    - real-time. fine-grained, efficient, easy to deployed, task-free for participants, affordable

# System Overview



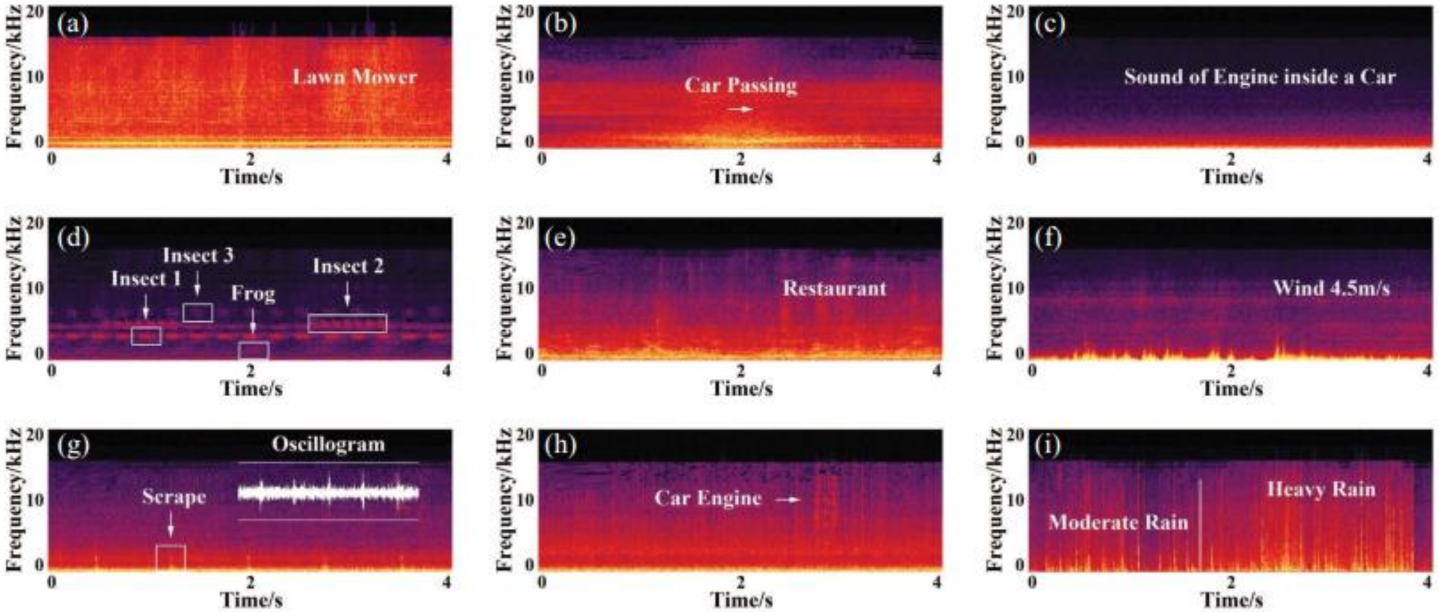


Fig. 2. Spectrograms of 4-s long audio segments which are crowdsourced from nine common scenes in our daily lives, including (a) lawn mower, (b) car passing, (c) sound of engine inside a car, (d) insects and a frog, (e) clamor of crowd in a restaurant, (f) wind, (g) scrapes of windshield wipers, (h) raindrops drumming against car window, and (i) raindrops drumming against umbrella.

Two challenges:

- background noises
- differences between different intensities of rain are small

# Part 1: Detection of Rain

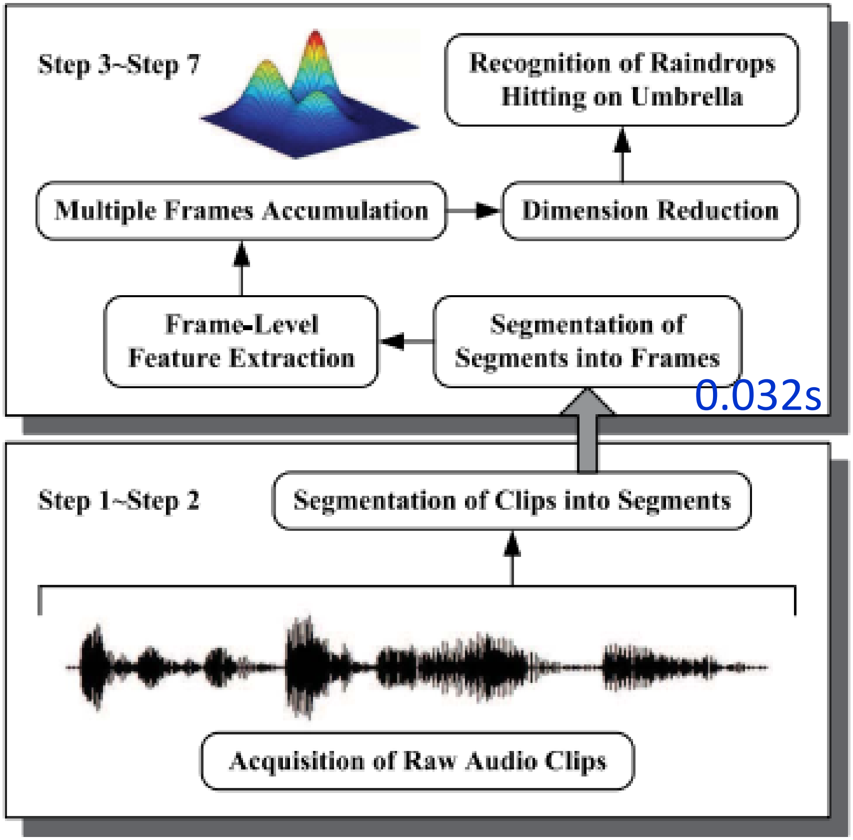
- Acquire raw audio segments of 1s
- extract features from power spectrum of each segment
    - min, med, max avg. amplitude
    - RMSE-L: measures the smooth degree of power spectrum curve
    - spectral similarity
    - spectral decrease
    - amplitude of middle freq
    - amplitude of cut-off freq
- Random Forest classifier

$$\text{RMSE-L}_i = \sqrt{\frac{1}{nl_i^p}\sum_{j=1}^{nl_i^p}\left(a_{ij}^p - \tilde{a}_{ij}^p\right)^2}.$$

$$\text{SS}_i = \frac{1}{nl_i^p}\sum_{j=1}^{nl_i^p}|\hat{a}_{ij}^p - \hat{a}_{i(nl_i+j)}^p|.$$

$$\text{SD}_i = \frac{1}{\sum_{j=2}^{n_i^p} a_{ij}^p}\sum_{j=2}^{n_i^p}\frac{a_{ij}^p - a_{i1}^p}{j-1}$$

# Part 2: Recognition of Raindrops Hitting on Umbrella



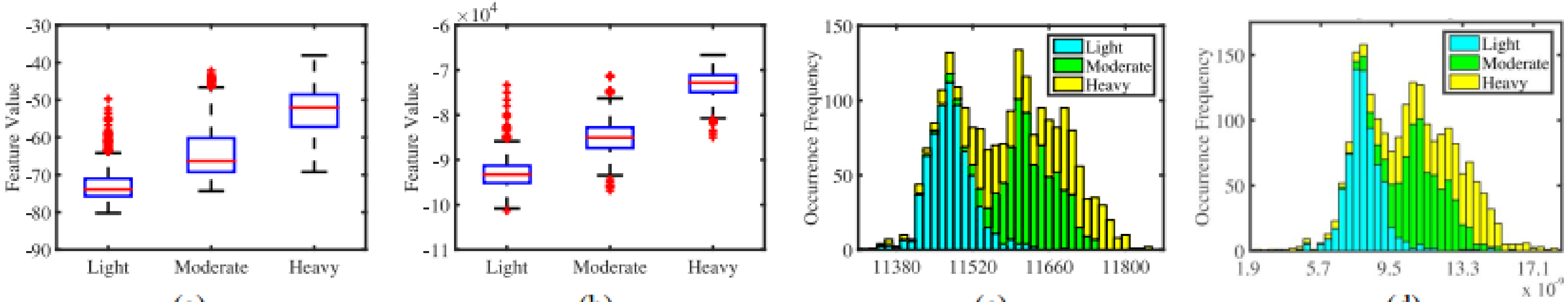
0.032s

Fig. 5. In recognition of raindrops hitting on umbrella. Computational process.

- extract features from each frames
  - avg. absolute amplitude
  - avg. zero-crossing rate
  - mel frequency cepstral coefficients
  - spectral centroid
- Multiple frames accumulation
  - use GMM to approximate the PDFs of feature vectors
  - MLE to solve for parameters
  - output 75-dimension features
- Dimensionality reduction
- Decision Tree Classifier

# Part 3: Measurement of Rain

- acquire segment level features in power spectrum
  - max amplitude: basic shape of every power spectrum curve
  - energy-low: signal energy in low frequency part
  - spectral roll-off: frequency below which 75% of the signal energy is contained
  - spectral slope: energy distribution at various freq
- Decision Tree Classifier

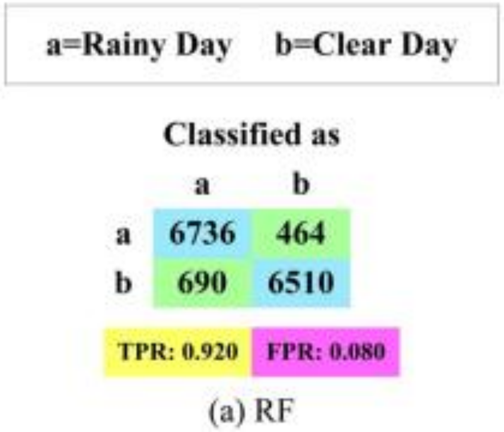


(a) (b) (c) (d)

Fig. 6. In measurement of rain. Box plots about MA, E-L, and histograms about SR-O, SS of 700 light, 700 moderate, and 700 heavy rain segments. (a) Max. Amplitude. (b) Energy-low. (c) SpectralRoll-off. (d) SpectralSlope-measurement.

# Evaluation Part 1

- 7200 rainy day and 7200 clear day segments
- 10-fold CV
- compare w/ existing rain detection method
    - 10 consecutive segments



a=Rainy Day b=Clear Day

Classified as

| | a | b |
|---|---|---|
| a | 6736 | 464 |
| b | 690 | 6510 |

TPR: 0.920 FPR: 0.080

(a) RF

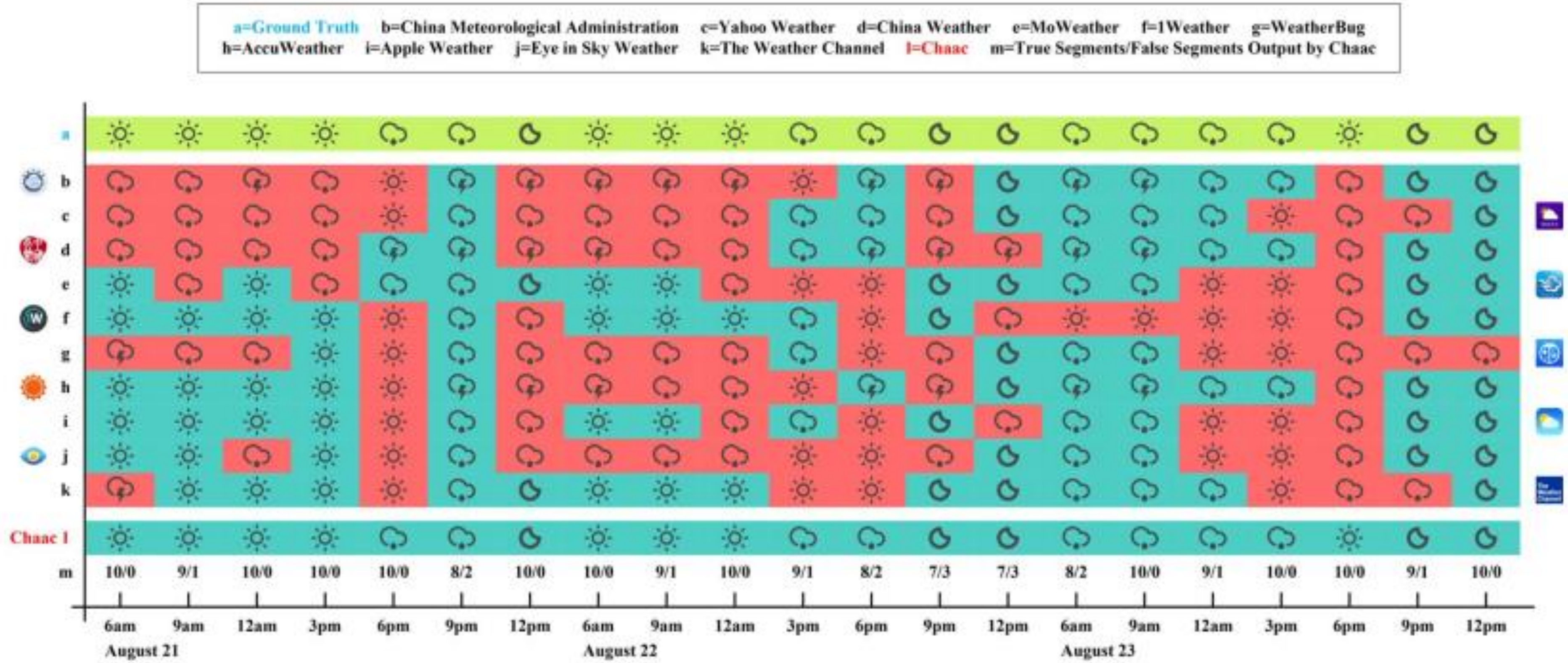


Fig. 8. In detection of rain. Comparison among the rain detection results output by Chaac, the rain information released by the CMA, and nine other popular weather apps, to show that Chaac works in real time.

# Evaluation Part 2

- 3600 segments of raindrops hitting umbrella and 3600 segments other activities in rainy day
- time consumption for steps
- Decision Tree + 10-fold CV
- compare different dimensionality reduction algorithms

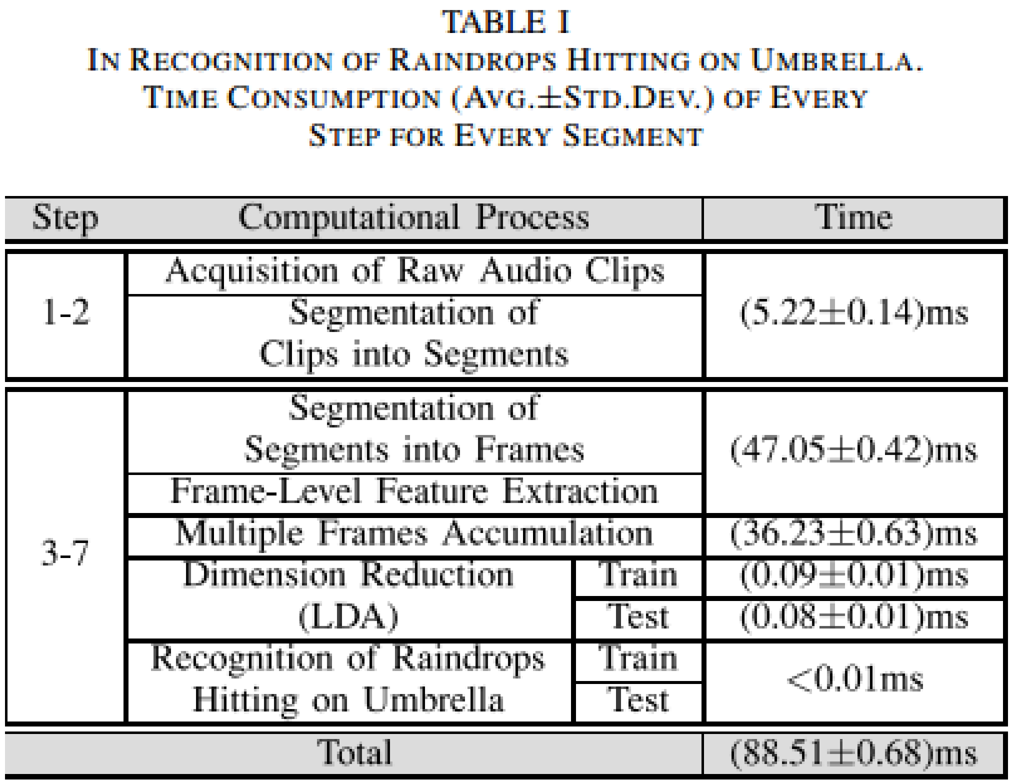
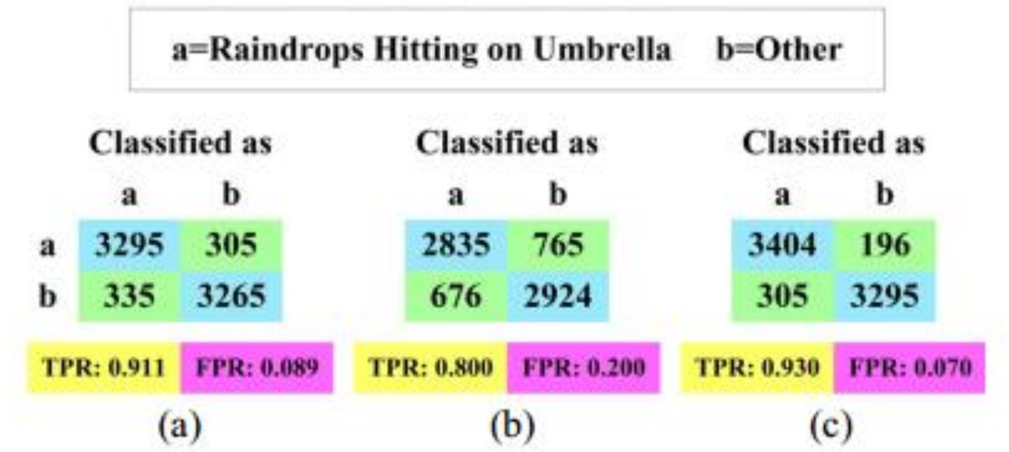


TABLE I
IN RECOGNITION OF RAINDROPS HITTING ON UMBRELLA.
TIME CONSUMPTION (AVG.±STD.DEV.) OF EVERY
STEP FOR EVERY SEGMENT

| Step | Computational Process | | Time |
|---|---|---|---|
| 1-2 | Acquisition of Raw Audio Clips | | (5.22±0.14)ms |
| | Segmentation of Clips into Segments | | |
| 3-7 | Segmentation of Segments into Frames | | (47.05±0.42)ms |
| | Frame-Level Feature Extraction | | |
| | Multiple Frames Accumulation | | (36.23±0.63)ms |
| | Dimension Reduction (LDA) | Train | (0.09±0.01)ms |
| | | Test | (0.08±0.01)ms |
| | Recognition of Raindrops Hitting on Umbrella | Train | <0.01ms |
| | | Test | |
| Total | | | (88.51±0.68)ms |



a=Raindrops Hitting on Umbrella b=Other

(a)

| | Classified as a | Classified as b |
|---|---|---|
| a | 3295 | 305 |
| b | 335 | 3265 |

TPR: 0.911 FPR: 0.089

(b)

| | Classified as a | Classified as b |
|---|---|---|
| a | 2835 | 765 |
| b | 676 | 2924 |

TPR: 0.800 FPR: 0.200

(c)

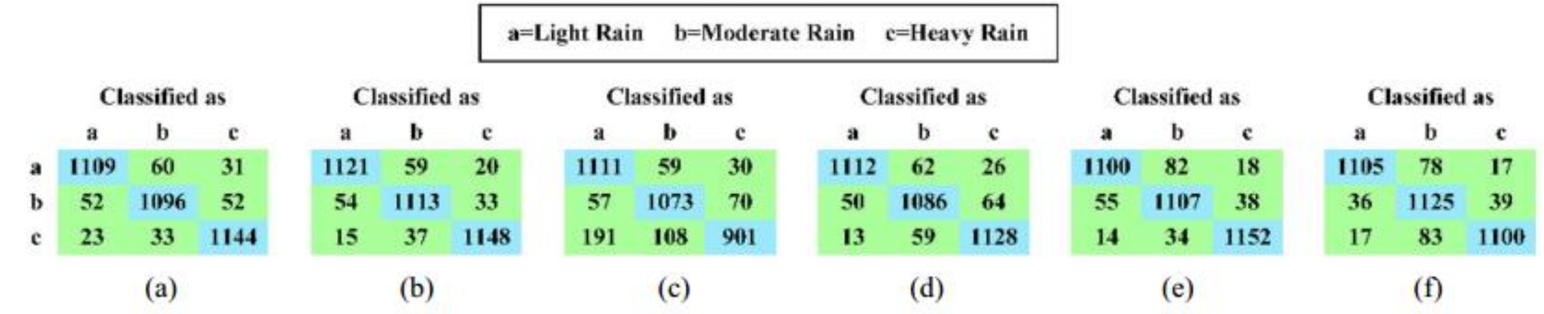
| | Classified as a | Classified as b |
|---|---|---|
| a | 3404 | 196 |
| b | 305 | 3295 |

TPR: 0.930 FPR: 0.070

Fig. 9. In recognition of raindrops hitting on umbrella. Confusion matrices, TPR, and FPR of different recognition algorithms. (a) DT. (b) PCA + DT. (c) LDA + DT.

# Evaluation Part 3

- 1200 light rain, 1200 moderate rain, 1200 heavy rain segments
- Decision Tree + 10-fold CV
- compare different ML models



a=Light Rain b=Moderate Rain c=Heavy Rain

(a) Classified as

| | a | b | c |
|---|---|---|---|
| a | 1109 | 60 | 31 |
| b | 52 | 1096 | 52 |
| c | 23 | 33 | 1144 |

(b) Classified as

| | a | b | c |
|---|---|---|---|
| a | 1121 | 59 | 20 |
| b | 54 | 1113 | 33 |
| c | 15 | 37 | 1148 |

(c) Classified as

| | a | b | c |
|---|---|---|---|
| a | 1111 | 59 | 30 |
| b | 57 | 1073 | 70 |
| c | 191 | 108 | 901 |

(d) Classified as

| | a | b | c |
|---|---|---|---|
| a | 1112 | 62 | 26 |
| b | 50 | 1086 | 64 |
| c | 13 | 59 | 1128 |

(e) Classified as

| | a | b | c |
|---|---|---|---|
| a | 1100 | 82 | 18 |
| b | 55 | 1107 | 38 |
| c | 14 | 34 | 1152 |

(f) Classified as

| | a | b | c |
|---|---|---|---|
| a | 1105 | 78 | 17 |
| b | 36 | 1125 | 39 |
| c | 17 | 83 | 1100 |

Fig. 10. In measurement of rain. Confusion matrices of different classifiers.

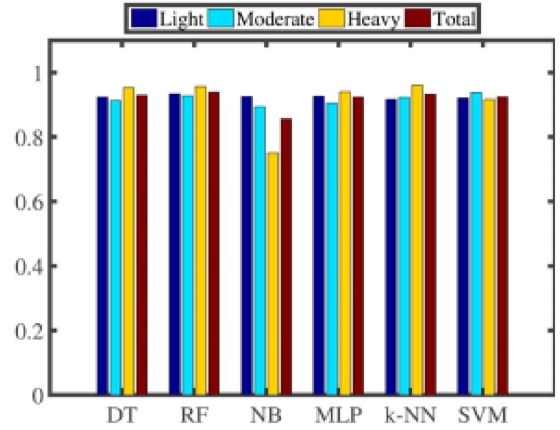
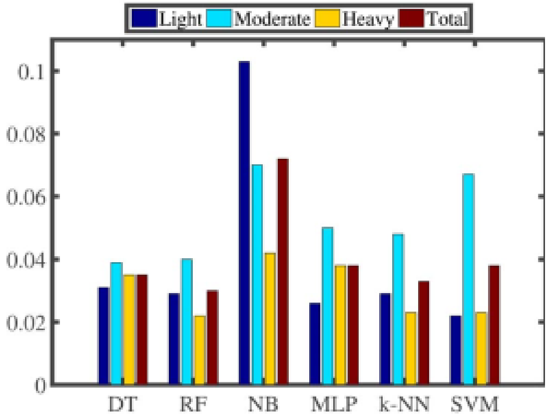
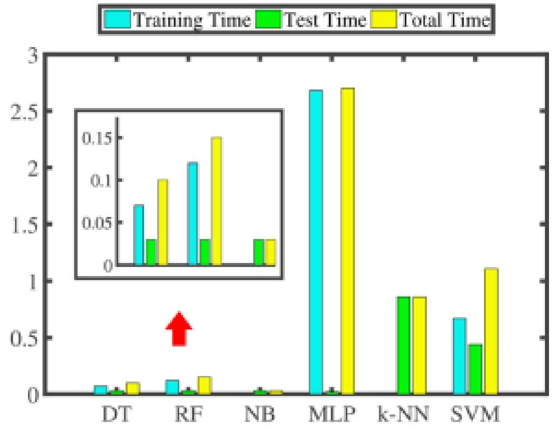


Fig. 11. In measurement of rain. TPR of different classifiers.



In measurement of rain. FPR of different classifiers.



In measurement of rain. Total time consumed of different classifiers.

# Real World Evaluation

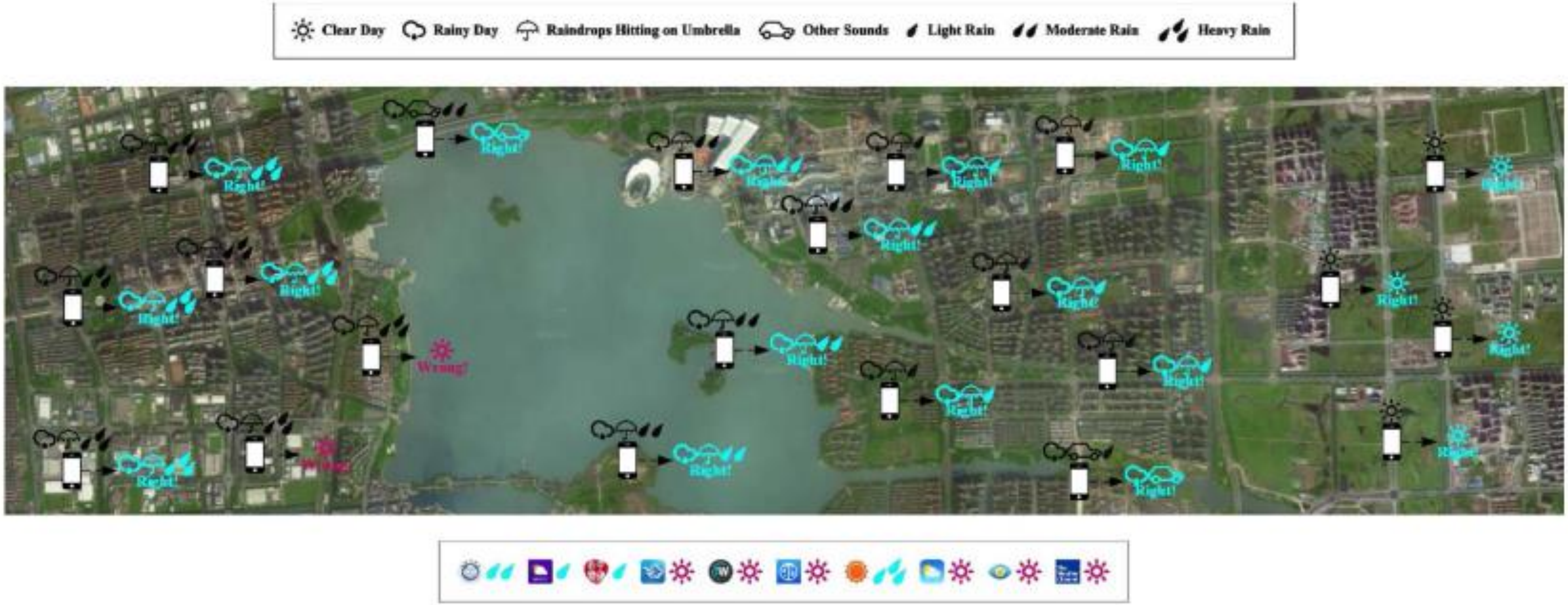


Fig. 14. Comparison among the rain detection and measurement results output by Chaac in real world, the rain information released by the CMA and nine other popular weather apps, to show that Chaac is real-time and fine-grained.

# Discussions and Conclusions

- Privacy Issue
  - outdoors environment
  - human voice is not of interest
  - processed locally and erased after use
  - user has complete control
- Motivation Issue
  - how to encourage people to share their data
- Remaining Meteorological Phenomena
  - snow, thunder, hail, hurricane, sandstorm
- Chaac: opportunistically crowdsourced audio clips from smartphones
  - real-time, fine-grained, economical

# Back Up