# InterGridNet: An Electric Network Frequency Approach for Audio Source Location Classification Using Convolutional Neural Networks

Christos Korgialas*, Ioannis Tsingalis*, Georgios Tzolopoulos*, and Constantine Kotropoulos

*Authors contributed equaly

Department of Informatics, Aristotle University of Thessaloniki, Greece

The Tenth International Conference on Advances in Signal, Image and Video Processing (SIGNAL 2025)

March 09 - March 13, Lisbon, Portugal

ARISTOTLE UNIVERSITY OF THESSALONIKI

IARIA

# Presenter's bio

**Constantine Kotropoulos** received the Diploma degree with honors in Electrical Engineering in 1988 and the PhD degree in Electrical & Computer Engineering in 1993, both from the Aristotle University of Thessaloniki. He is a Full Professor in the Department of Informatics at the Aristotle University of Thessaloniki. He was a visiting research scholar in the Department of Electrical and Computer Engineering at the University of Delaware, USA during the academic year 2008-2009. He also conducted research in the Signal Processing Laboratory at Tampere University of Technology, Finland during the summer of 1993. He has co-authored 71 journal papers and 226 conference papers and contributed 9 chapters to edited books in his areas of expertise. He is co-editor of the book Nonlinear Model-Based Image/Video Processing and Analysis (J. Wiley and Sons, 2001).
His research interests include forensics, audio, speech and language processing, signal processing, pattern recognition, multimedia information retrieval, biometrics, and forensics.
Prof. Kotropoulos was a scholar of the State Scholarship Foundation of Greece and the Bodossaki Foundation. He is a senior member of the IEEE and a member of EURASIP, IAPR, and the Technical Chamber of Greece. He was a Senior Area Editor of the IEEE Signal Processing Letters. He has been a member of the Editorial Board of the journals Advances in Multimedia, International Scholar Research Notices, Computer Methods in Biomechanics & Biomedical Engineering: Imaging & Visualization, Artificial Intelligence Review, MDPI Imaging, MDPI Signals, and MDPI Methods and Protocols. Prof. Kotropoulos served as Track Chair for Signal Processing in the 6th Int. Symposium on Communications, Control, and Signal Processing, Athens, 2014; Program Co-Chair of the 4th Int. Workshop on Biometrics and Forensics, Limassol, Cyprus, 2016; Technical Program Chair of the XXV European Signal Processing Conf., Kos, Greece, 2017; Technical Program Chair of the 5th IEEE Global Conf. Signal and Information Processing, Montreal, Canada, 2017; General Chair of the 2022 IEEE 14th Image, Video, and Multidimensional Signal Processing Workshop, Nafplio, Greece; Technical Program Chair of the 2023 IEEE International Conf. on Acoustics, Speech, and Signal Processing, Rhodes, Greece.

# Outline

- **Electric Network Frequency (ENF)** fluctuates around 50 Hz (Europe) and 60 Hz (US/Canada) due to power grid load variations. These fluctuations are embedded into multimedia recordings and can be extracted for forensic applications.
- **Grid Localization** is categorized into inter-grid (identifying the power grid where a recording was made) and intra-grid (precisely locating the recording within a specific grid).
- **ENF Variability and Challenges** arise from factors such as local load changes, generator failures, and power line switching. City-wide power consumption patterns and grid topology further impact ENF characteristics, affecting localization accuracy.

# Introduction

## Contributions

- **InterGridNet Framework:** A deep learning framework for geolocation classification using ENF signatures, optimized with Neural Architecture Search (NAS).
- **Shallow RawNet Model:** Utilizes a lightweight RawNet[1] to minimize parameters while maintaining performance, making it more efficient for inter-grid localization.
- **End-to-End Processing Pipeline:** Introduces a comprehensive approach, including preprocessing, ENF signal isolation, feature extraction via residual layers, and classification using softmax activation.
- **State-of-the-Art Performance:** Evaluated on the SP Cup 2016 dataset[2], achieving high accuracy across nine power grids and outperforming existing geolocation classification methods.

[1] J.-W. Jung, H.-S. Heo, J.-H. Kim, H.-J. Shim, and H.-J. Yu. "RawNet: Advanced end-to-end deep neural network using raw waveforms for text-independent speaker verification". *arXiv preprint arXiv:1904.08104* (2019)

[2] M. Wu, A. Hajj-Ahmad, M. Kirchner, Y. Ren, C. Zhang, and P. Campisi. "Location Signatures That You Don't See: Highlights from the IEEE Signal Processing Cup Student Competition". *IEEE Signal Processing Magazine* 33.5 (2016), pp. 149–156

- **Dataset Overview**: The SP Cup 2016 dataset is used, consisting of audio and power recordings from nine power grids (**A** to **I**).
- **Grid Characteristics**:
  - Grids **A**, **C**, and **I** are characterized by 60Hz nominal ENF.
  - Other grids exhibit ENF around 50Hz.
- **Recording Types**:
  - **Power Recordings**: Captured from the power mains (stronger ENF traces).
  - **Audio Recordings**: Captured near electrical devices (weaker ENF traces).
- **Dataset Splits**:
  - **Training Set**: For model development and training.
  - **Practice Set**: For model validation.
  - **Testing Set**: Contains unseen data for performance evaluation.

# Methodology

## Data Preprocessing

- **Augmentation**:
  - Testing set augmented with 100 samples (40 Audio, 60 Power) and 10 "None" (**N**) samples from other grids.
- **Categorization**:
  - Recordings classified into four groups: `audio50`, `audio60`, `power50`, and `power60`.
  - Automatic categorization based on spectral characteristics during testing.
- **Spectral Classification**:
  1. Calculate average spectrogram magnitudes for 50Hz and 60Hz harmonics.
  2. Ignore the weakest harmonic; use the strongest two for classification.
- **Preprocessing**:
  - Bandpass filter isolates ENF signal within a 2Hz range (e.g., `audio60`: 59Hz–61Hz).
  - Segment into 16-second frames with 50% overlap, and normalize to $[-1, 1]$.

# Methodology

## InterGridNet Framework

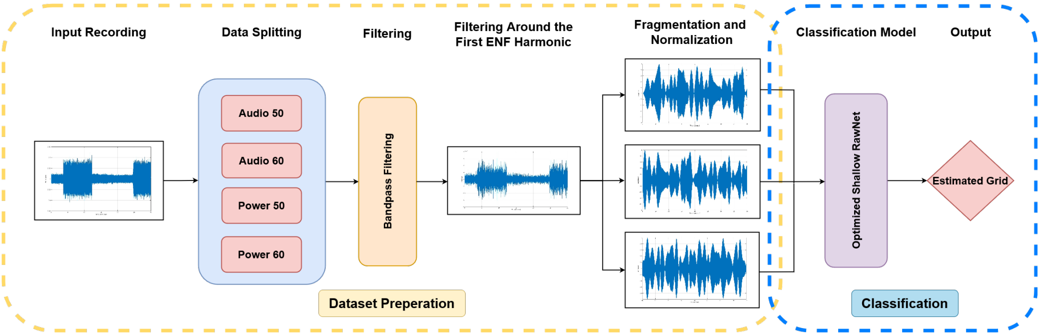



Figure 1: Flowchart of the proposed InterGridNet framework.

# Methodology

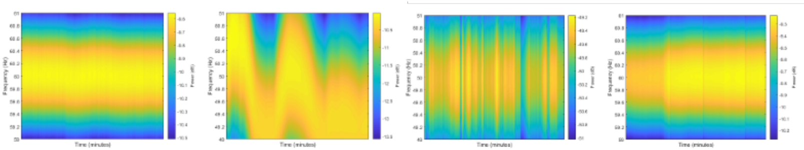
## Classification Problem

**Spectral Variation Across Grids:**

- Spectrograms reveal distinct ENF characteristics per grid.
- Nominal ENF: 60Hz for Grids **A**, **C**, and **I**; 50Hz for the remaining grids.

**Classification Problem:**

- Samples: 16-second frames categorized by grid.
- Defined as $G^{50}_{\mathrm{Audio}}$, $G^{60}_{\mathrm{Audio}}$, $G^{50}_{\mathrm{Power}}$, $G^{60}_{\mathrm{Power}}$.
- Number of classes: $n = 3$ for 60Hz groups, $n = 6$ for 50Hz groups.



Grid A    Grid B    Grid D    Grid I

Figure 2: Spectrograms focused on the nominal ENF value for different grids.

## Neural Network for Classification:

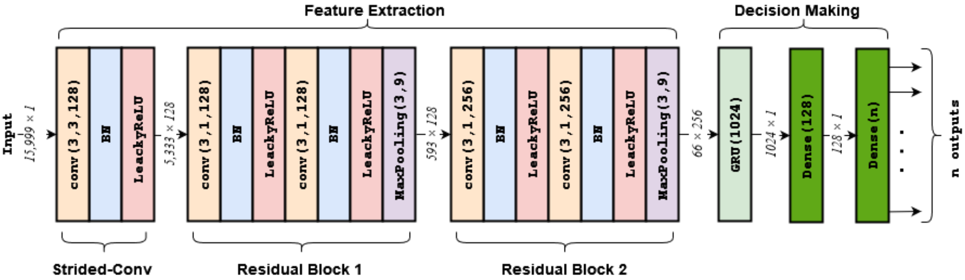



Figure 3: Architecture of the optimized shallow RawNet model.

- Implemented a shallow RawNet architecture optimized using NAS[3].

Table 1: Optimized hyperparameters for the shallow RawNet model.

| | $G^{50}_{\text{Audio}}$ | $G^{60}_{\text{Audio}}$ | $G^{50}_{\text{Power}}$ | $G^{60}_{\text{Power}}$ |
|---|---|---|---|---|
| Learning Rate | $6.5 \times 10^{-4}$ | $7 \times 10^{-4}$ | $1.1 \times 10^{-3}$ | $9.7 \times 10^{-4}$ |
| $\beta_1$ | 0.96 | 0.97 | 0.98 | 0.98 |
| $\beta_2$ | 0.998 | 0.998 | 0.992 | 0.993 |

[3] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama. "Optuna: A Next-Generation Hyperparameter Optimization Framework". In: *Proceedings 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 2019, pp. 2623–2631

**Hyperparameter Optimization:**

- Learning rate and Adam optimizer parameters tuned using NAS.
- Optimized values shown in Table 1.

**Final Classification Decision:**

- Classification based on entropy threshold:

$$-\sum_{i=1}^{n} p_i(x) \log_2 p_i(x) < \alpha_1 \cdot \log_2(n), \tag{1}$$

- If not satisfied, sample labeled as **N**.
- Majority voting mechanism ensures robustness:
  - Final estimate valid if it appears in at least $\alpha_2$ of frames.
  - Otherwise classified as **N**.
- Parameters: $\alpha_1 = 0.8$, $\alpha_2 = 0.75$.

- **Model Validation:** 50 samples from SP Cup 2016 dataset.
- **Validation Accuracy:**
  - **Overall:** 90% accuracy.
  - **Best:** Power recordings (96.67%).

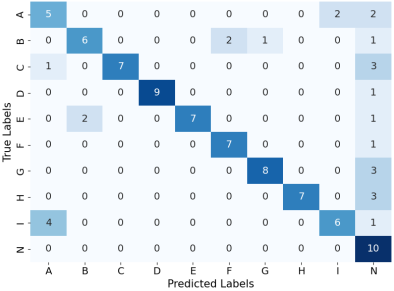
Table 2: InterGridNet validation accuracy.

| **Type** | **A** | **B** | **C** | **D** | **E** | **F** | **G** | **H** | **I** | **N** | **Overall** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Audio | 80% | 100% | 100% | 100% | 80% | 100% | 80% | 80% | 100% | 100% | 80% |
| Power | 100% | 100% | 100% | 100% | 100% | 100% | 80% | 100% | 100% | 100% | 96.67% |
| All | 80% | 100% | 100% | 100% | 80% | 100% | 60% | 80% | 100% | 100% | 90% |

**Confusion Matrices:**

- **Filtered Data:** 92% accuracy (Figure 4(a)).
- **Non-Filtered Data:** 72% accuracy (Figure 4(b)).


(a) Filtered Data (92% accuracy)




(b) Non-Filtered Data (72% accuracy)

Figure 4: Confusion matrices.

# Discussion

## Comparison with Other Methods

- **Testing Accuracy Comparison:** The proposed **InterGridNet** framework outperforms previous methods with a 92% accuracy.
- **Other Methods:** Comparison with existing methods shows **InterGridNet** offers a higher accuracy for grid localization.

Table 3: Testing Accuracies (%) in SP Cup 2016 Dataset.

| **Method** | **Characteristic** | **Accuracy** |
|---|---|---|
| SVM[4] | One-vs-one classification | 86% |
| SVM[5] | Multi-class classification | 77% |
| SVM[6] | Multi-class classification | 88% |
| Random Forest, SVM, AdaBoost[7] | Ensemble method | 88% |
| Binary SVM[8] | Binary classification | 87% |
| Multi-Harmonic Histogram Comparison[9] | Frequency domain analysis | 88% |
| **InterGridNet (Ours)** | Shallow RawNet | **92%** |

4 A. Triantafyllopoulos et al. *Exploring Power Signatures for Location Forensics of Media Recordings*. Tech. rep. Signal Processing Cup. University of Patras, Greece, 2016

5 R. Ohib, S. Y. Arnob, R. Arefin, M. Amin, and T. Reza. "ENF Based Grid Classification System: Identifying the Region of Origin of Digital Recordings". *Criterion* 3.4 (2017), p. 5

6 H. Zhou et al. *Geographic location estimation from ENF signals with high accuracy*. Tech. rep. Signal Processing Cup. University of Science and Technology of China, 2016

7 M. El Helou, A. W. Turkmani, R. Chanouha, and S. Charbaji. "A Novel ENF extraction approach for region-of-recording identification of media recordings". *Forensic Science International* 155.2-3 (2005), p. 165

8 D. Despotović et al. *Exploring Power Signatures for Location Forensics of Media Recordings*. Tech. rep. Signal Processing Cup. University of Novi Sad, Serbia, 2016

9 C. Chow et al. *Multi-harmonic histogram comparison*. Tech. rep. Signal Processing Cup. Purdue University, 2016

# Conclusions

- **InterGridNet**: A novel framework for geolocating audio recordings across power grids, optimized using NAS.
- Built on RawNet's architecture, **InterGridNet** uses a shallow version of RawNet to address complex inter-grid localization.
- Achieved **92% accuracy** on the SP Cup 2016 dataset, marking the first DNN-based approach with preprocessing methods.
- **Future work** includes:
  - Adoption of transformer architectures for improved grid location classification.
  - Integration of explainable AI (xAI) to enhance transparency and decision-making.

# Github Code

**The code for the proposed framework can be found at:**

`https://github.com/ckorgial/InterGridNet`

or by scanning the QR code:

# Acknowledgments

This research was supported by the Hellenic Foundation for Research and Innovation (HFRI) under the "2nd Call for HFRI Research Projects to support Faculty Members & Researchers" (Project Number: 3888).

# Thank You!

Thank you very much for your attention.
Q & A?
**Email** costas@csd.auth.gr