Non-Intrusive Residential Electric Vehicle Detection Using Positive-Unlabeled Contrastive Learning

Zahrasadat Sajjadifar[1], Jan Yperman[2], Oscar Mauricio Agudelo[1], Antonio Hernandez Espin[2], Thijs Becker[2], Koen Vanthournout[2], Bart De Moor[1]

[1]STADIUS Center for Dynamical Systems, Signal Processing, and Data Analytics, Department of Electrical Engineering (ESAT), KU Leuven, Leuven, Belgium

[2]Water & Energy Transition (WET), VITO/EnergyVille, Genk, Belgium

{zahrasadat.sajjadifar, mauricio.agudelo, bart.demoor}@kuleuven.be

{jan.yperman, antonio.hernandezespin, thijs.becker, koen.vanthournout}@vito.be

Abstract

Detecting electric vehicle (EV) charging on residential connections is essential for distribution system operators (DSOs) to manage grid load, forecast demand, and plan infrastructure upgrades. However, the widespread availability of labeled data for such detection remains limited, especially in non-intrusive settings. Notably, while there is often a reliable positive label for EV presence through EV registration, the negative labels tend to be unreliable due to the presence of non-registered EVs. This paper addresses the problem of EV charging detection as a positive and unlabeled (PU) contrastive learning task, where only a subset of the measurements is positively identified based on the households registered with EV, and the rest of the data remains unlabeled. Based on contrastive learning, we learn representations that pull together examples likely to contain EV charging signatures and push apart background load patterns and non-EV signatures, while explicitly accounting for the label uncertainty inherent in PU data. We propose an approach that works with raw, aggregated electricity load data at the household level, without relying on intrusive metering or extensive manual labeling. Our primary dataset consists of quarter-hourly household electricity consumption data provided by Fluvius, the distribution system operator (DSO) in Flanders, Belgium, and we additionally validate our method using other publicly available open datasets. Our results highlight the feasibility of scalable EV detection with minimal supervision, offering critical observability for DSOs aiming to monitor EV adoption and manage localized grid impacts.

1 Introduction

The rapid adoption of electric vehicles (EVs) is reshaping electricity consumption patterns in residential areas, introducing new challenges for distribution system operators (DSOs). Unlike traditional household loads, EV charging sessions can be substantial, sporadic, and time-dependent, creating considerable stress on low-voltage distribution networks (Becker et al. 2025). Accurately detecting residential with EV charging events is therefore critical for DSOs to ensure grid stability, optimize demand response, and plan future infrastructure investments. However, identifying households with EV charging from aggregated residential electricity consumption data is a non-trivial task. In many real-world scenarios, a separate submeter for EV is unavailable, and therefore ground-truth labels for EV charging sessions are missing. This creates a significant barrier for the application of traditional supervised machine learning techniques, which typically require large amounts of labeled data to achieve reliable performance. To address this challenge, we propose framing the EV detection problem as a positive and unlabeled (PU) learning task, which reflects the realistic constraints of partial labeling: some households with EV are known through EV registration (positive samples), but most of the households are unlabeled, and due to the non-registered households with EV, this subset contains a mix of EV and non-EV consumption. This formulation aligns well with the nature of residential energy data and allows for the application of semi-supervised learning strategies.

In this work, we present a non-erosive, end-to-end clustering pipeline for identifying households with EV charging activity from load profiles. Unlike methods that rely on extensive preprocessing or signal transformation, which can distort the original load signature, our approach preserves the temporal structure of the data and seeks to uncover charging patterns directly from the raw electricity consumption. By using representation learning and clustering techniques, we aim to identify consistent EV charging behaviors in the absence of explicit labels.

Our contributions are: (1) we frame EV detection as a PU-contrastive learning problem and analyze its implications for residential energy monitoring, (2) we introduce an end-to-end pipeline capable of identifying households with and without EV from aggregated load data, (3) we evaluate our approach on multiple datasets and show that it can learn meaningful representations for identifying households with and without EV, offering a scalable solution for DSOs.

We review related work and describe our methodology, experiments, and results in detail in the following sections.

2 Related Works

Prior research on detecting EVs from residential electricity consumption can be broadly divided into two categories: (1) EV charging sequence or event detection and (2) household-level EV ownership identification. Event detection approaches aim to locate and characterize individual charging sessions from aggregated load profiles. For example, (Vavouris et al. 2022) used Random Forests with load reconstruction and deep neural networks with GAN-based post-processing for generalizable EV charging disaggregation. (Wang et al. 2022) proposed EVSense, a deep learning framework for robust large-scale detection trained with labeled EV charging events. (Yin et al. 2023) presented a multi-kernel convolutional neural network (CNN) for substation-level EV load disaggregation, requiring detailed ground-truth for EV power sequences. (Martin, Ke, and Wang 2023) applied sliding-window feature extraction with classical machine learning models like XGBoost and Random Forest for feeder-level event inference using labeled EV charging activity. In the unsupervised and training-free space, (Zhang et al. 2014) detected EV events from low-sampling-rate data without any ground-truth, using a rule-based algorithm that combines signal thresholding and filtering to isolate EV charging patterns from household loads. While (Ghaffar et al. 2022) and (Criado-Ramón et al. 2024) used clustering and graph-based techniques for appliance disaggregation, these methods can also be adapted to EV detection. Household-level identification methods instead determine whether a home owns an EV without localizing specific charging events. (Neubert et al. 2022) employed a supervised CNN–MLP classifier on smart meter data to detect EV presence. (Hoffmann and Fesche 2019) used convolutional and recurrent models to classify EV households from hourly data. Similarly, (Aly et al. 2024) introduced RES-EV, a residual-based model for EV household detection under high AC loads. Several of these studies rely on datasets in which EV charging is submetered as a separate load, providing full ground-truth labels for EV events.

Our work also addresses the problem of identifying households with EV but departs from prior supervised classification frameworks. Instead of directly training a classifier, we learn discriminative representations of EV and non-EV households from positive and unlabeled data using contrastive learning, followed by clustering to distinguish household types. Unlike previous studies using synthetic data or explicit EV submetering, our approach leverages real-world utility data with limited but reliable positive labels obtained from EV registration process, ensuring practical relevance and robustness.

3 Preliminary

In this section, we briefly define the two core components of the proposed framework: Positive-Unlabeled (PU) learning and contrastive learning. These paradigms enable learning meaningful representations from data where only partial supervision is available.

3.1 Positive-Unlabeled Learning

Positive-Unlabeled (PU) learning is a weakly supervised learning setting in which only positive examples and unlabeled data are available during training. Unlike standard supervised learning, where both positive and negative labels are provided, PU learning must account for the fact that the unlabeled set contains a mixture of both positive and negative examples (Bekker and Davis 2020). This setup is particularly relevant in real-world scenarios where obtaining negative labels is costly, ambiguous, or impossible.

Formally, let $X = \{x_1, x_2, \ldots, x_n\}$ denote a dataset of instances, where a small subset $X_P \subset X$ are labeled as positives, and the remainder $X_U = X \setminus X_P$ are unlabeled. The goal is to learn a representation that distinguishes latent positive instances from negative instances in U using information from P. PU learning techniques generally fall into two categories: (1) those that estimate the class prior and reweight examples accordingly (e.g., unbiased risk estimators), and (2) those that use heuristics or iterative re-labeling strategies to identify likely negatives or pseudo-label the unlabeled data.

3.2 Contrastive Learning

Contrastive learning is a self-supervised representation learning that learns to distinguish between similar and dissimilar data points by optimizing a contrastive loss. The core idea is to bring embeddings of similar (positive) pairs closer together in the embedding space while pushing apart embeddings of dissimilar (negative) pairs. Formally, given an input x, an encoder function f maps x to a latent representation $h = f(x)$. This representation is then passed through a projection head g to obtain latent embedding $z = g(h)$, which is used to compute the contrastive loss.

Typically, positive pairs are created through multiple stochastic augmentations of the same input, while negative pairs are formed from other instances in the augmented batch. A common loss function used in contrastive learning is the InfoNCE loss (van den Oord, Li, and Vinyals 2019):

$$\mathcal{L}_{\text{InfoNCE}} = \sum_{i \in \mathbf{I}} -\log \frac{\exp(z_i, z_{a(i)}/\tau)}{Z_i} \tag{1}$$

$$Z_i = \sum_{k \in \mathbf{I} \setminus \{i\}} \exp(z_i \cdot z_k/\tau) \tag{2}$$

where τ is a temperature hyperparameter that controls the sharpness of the similarity distribution. The set $\mathbf{I} = \{1, \ldots, 2b\}$, with b as batch-size, indexes all samples in the two-view augmented batch $\tilde{D}_t$, and $a(i)$ denotes the index of the positive (augmented) pair corresponding to sample i. The denominator Z_i sums over one positive and $2(b-1)$ negatives, i.e., all other samples in the batch. This loss encourages the model to correctly identify the positive example among a set of negatives given an anchor.

SimCLR (Chen et al. 2020) is a popular framework for visual representation learning. It uses a CNN backbone, typically a ResNet, to extract latent representation, followed by a projection head that maps representation into the embedding space where the contrastive loss is applied. The learning objective is the InfoNCE loss, which encourages the network to produce semantically meaningful representation suitable for downstream visual tasks such as classification.

3.3 PU-Contrastive Learning

In PU settings, constructing reliable negative pairs is challenging because the unlabeled data may contain hidden positives. To address this, recent approaches in PU-contrastive

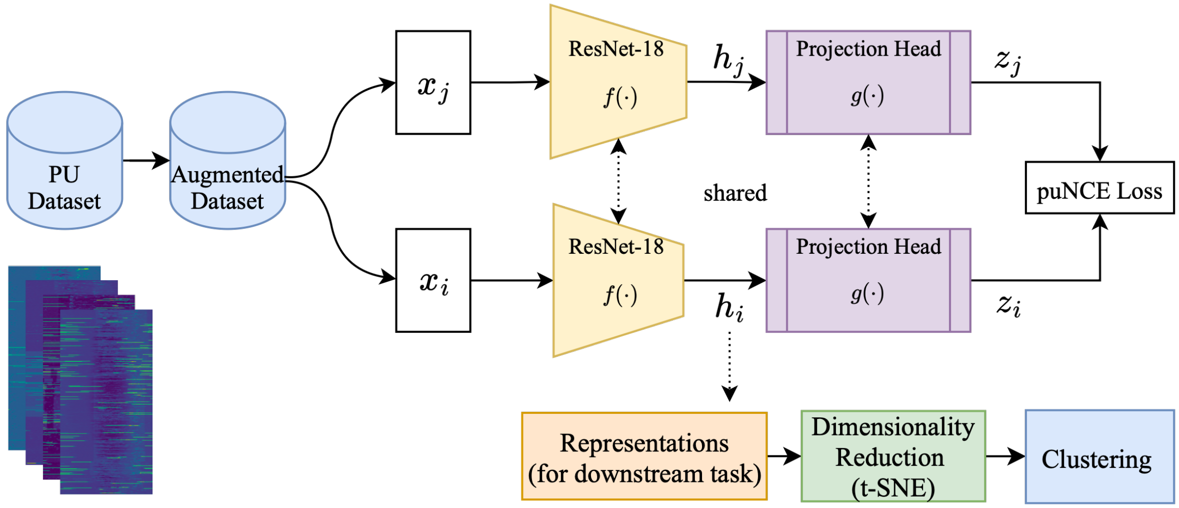

Figure 1: Overview of the proposed Positive-Unlabeled (PU) contrastive learning pipeline for identifying households with EV. Raw household electricity consumption time series are transformed into two-dimensional heatmap representation and augmented to create multi-view batches. Each image is encoded by a ResNet-18 backbone $f(\cdot)$ to produce representations h, which are then passed through a projection head $g(\cdot)$ to obtain embeddings z for optimizing the puNCE loss. After training, the projection head is discarded and the encoder outputs h are used as representations for downstream task. These representations are further processed with dimensionality reduction (t-SNE) and clustering (e.g., k-means, spectral clustering) to infer EV presence.

learning (Acharya et al. 2024) proposed adjusted loss formulations. In particular, puNCE (positive unlabeled InfoNCE) extends the self-supervised InfoNCE loss to the PU learning setting by treating each unlabeled sample as positive with probability π and negative with probability $1-\pi$. This formulation reduces the penalty for treating actual positive samples in the unlabeled set as negative, thus preserving representation quality. For a labeled anchor $x_i \in X_P$, puNCE pulls together embeddings of all labeled samples in the augmented batch $\tilde{D}_t$, unlike InfoNCE which considers only a single positive. Let $\mathbf{P}$ and $\mathbf{U}$ denote the labeled and unlabeled indices where $\mathbf{I} = \mathbf{P} \cup \mathbf{U}$. The empirical risk for labeled samples is:

$$\ell_P = -\sum_{i \in \mathbf{P}} \frac{1}{|\mathbf{P}|-1} \sum_{j \in \mathbf{P} \setminus \{i\}} \log \frac{\exp(z_i \cdot z_j / \tau)}{Z_i} \quad (3)$$

For unlabeled anchors $x_i \in X_U$, puNCE considers them positive with probability π and negative with $1-\pi$, weighting the contributions accordingly:

$$\ell_U = -\sum_{i \in \mathbf{U}} \Bigg[\pi \frac{1}{|\mathbf{P}|+1} \sum_{j \in \mathbf{P} \cup \{a(i)\}} \log \frac{\exp(z_i \cdot z_j / \tau)}{Z_i} + (1-\pi) \log \frac{\exp(z_i \cdot z_{a(i)} / \tau)}{Z_i} \Bigg] \quad (4)$$

Then, the total empirical loss is given by: $\mathcal{L}_{\text{puNCE}} = 1/2b\,[\ell_P + \ell_U]$. Essentially, puNCE assigns unit weight to labeled samples, while unlabeled samples are duplicated and weighted by π and $1-\pi$ for positive and negative contributions, respectively.

4 Methodology

Our objective is to identify households with EV under partial supervision, where only a small subset of households are labeled as EV users through EV registration process and the remaining households are unlabeled. To address this, we design a clustering pipeline based on Positive-Unlabeled (PU) contrastive learning that operates on heatmap representation of household electricity consumption time series. The overall workflow of pipeline is illustrated in Figure 1, outlining the process from raw consumption data to the learned representation used for clustering.

The approach treats labeled EV households as anchors and learns latent embeddings that distinguish them from unlabeled households, while explicitly incorporating prior knowledge about the uncertainty inherent in the unlabeled set. By leveraging both contrastive objectives and PU learning principles, the model captures consumption patterns indicative of EV charging behavior even when explicit negative labels are absent. Finally, the learned representations are clustered to infer EV presence across all households.

4.1 Time Series Heatmap Representation

To capture meaningful consumption patterns over time, we represent energy usage of each household as a two-dimensional heatmap. Specifically, we reshape the time series into a matrix where columns correspond to time intervals within a day (e.g., hours of day) and rows represent sequential days. Each element in the heatmap encodes energy consumption at a specific time and day. This format enables the model to capture both short-term (daily) and long-term

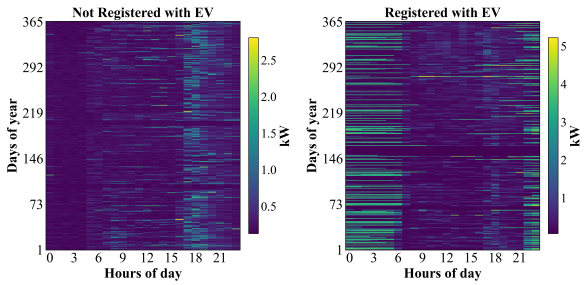

Figure 2: Time series heatmap representations of two households from Fluvius dataset. The household with an registered EV (right) shows distinct recurring high-consumption patterns, while the household with no registered EV does not. This illustrates how heatmap patterns can serve as a strong visual indicator of EV charging behavior.

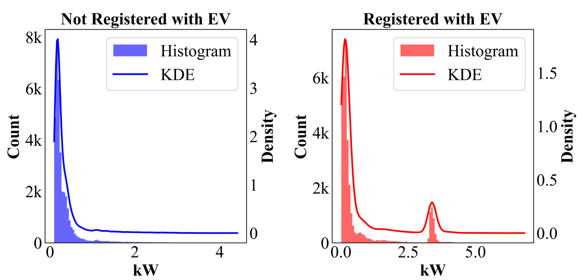

Figure 3: Histogram and KDE of aggregated load for two households from Fluvius dataset. The household with registered EV (right) shows a pronounced peak in the typical EV charging range, unlike the non-EV household (left), illustrating that detecting peaks in KDE can be an informative feature to identify households with EV.

(weekly or monthly) temporal patterns that are indicative of EV charging behavior, such as repeated high-consumption events during evening hours.

The heatmap representation allows us to treat the problem as a visual clustering task, where we can leverage CNNs to learn spatially-localized patterns. Compared to raw time series or hand-crafted features, this approach preserves the temporal structure of the data while enabling powerful and scalable feature extraction through modern deep learning techniques. Figure 2 illustrates two exemplary households, one with registered EV and one without. As can be seen, even to a human expert, the household with an EV exhibits distinct recurring evening consumption patterns, while the non-EV household does not, demonstrating the discriminative power of the heatmap representation.

4.2 Model Architecture

Inspired by the SimCLR framework (Chen et al. 2020), we adopt a contrastive learning approach tailored to the PU setting, as described in Section 3. While SimCLR is designed for three-channel images, our inputs are two-dimensional heatmap representations of household electricity consumption time series. To accommodate this, we modify the first convolutional layer of a ResNet-18 backbone to accept single-channel heatmaps. The encoder maps each heatmap into a compact latent representation of size 512 for standard ResNet-18, which is then passed through a projection head consisting of a two-layer multilayer perceptron (MLP) with ReLU activation that maps the latent representations to latent embedding where contrastive loss is applied.

We employ the puNCE loss (Acharya et al. 2024) by weighting unlabeled samples according to their estimated probability of being positive. Under the SCAR (Selected Completely At Random) assumption, we set the class prior $\pi = 0.15$ based on domain expertise.

4.3 Data Augmentation Strategy

Multi-view batches are generated through data augmentations inspired by the empirical survey in (Iwana and Uchida 2021) including random scaling, additive noise, and temporal masking which randomly masks several days of consumption data in the heatmap by replacing the values with zeros. This augmentation improves robustness to varying data availability during validation or deployment, as the model can generalize from partial time series to an extended one-year span.

4.4 Clustering and Label Inference

After contrastive training converges, we obtain a fixed-length representation for each household. These representations are designed to capture differences between EV and non-EV consumption patterns. We choose clustering as a downstream task to assess whether the learned features in the representation space are truly discriminative of EV signatures. To facilitate clustering and mitigate the curse of dimensionality, we first apply t-SNE (Maaten and Hinton 2008) to reduce the high-dimensional representation to a two-dimensional space that preserves neighborhood relationships. We select t-SNE because it preserves local neighborhoods, captures non-linear relationships in the data, and often produces well-separated clusters in two-dimensional space, making it suitable for clustering and visualization. The resulting components are then normalized to have zero mean and unit variance, ensuring balanced feature scaling before clustering. We evaluate multiple clustering methods on the normalized t-SNE components, including k-means and spectral clustering, both configured with $k = 2$ (number of clusters). Cluster-to-class assignments are determined via majority voting: the cluster containing the highest proportion of positive labels is designated as the EV-positive class, while the other is treated as the negative class.

This approach remains non-parametric and leverages the learned representation space to separate EV consumption patterns from background load is a positive and unlabeled setting. The result is an end-to-end clustering pipeline that requires neither event detection nor signal disaggregation, relying instead on the aggregated consumption signature over time.

Dataset	Metric	#Instances [Pos / (Neg/Unlabeled)]	k-means Clustering	Spectral Clustering	Baseline
Fluvius (Train)	Recall	2455 / 3904	0.81	**0.86**	0.67
Fluvius (Test)	Recall	614 / 976	0.82	**0.84**	0.67
Irish_synthetic	F1-score	1919 / 1919	0.91	**0.99**	0.92
UK	F1-score	99 / 193	0.68	**0.68**	0.52

Table 1: Clustering performance on PU–contrastive representation. Feature representations are learned via PU-contrastive learning and reduced to two dimensions using t-SNE, k-means and spectral clustering are applied. Recall is reported for the PU-labeled Fluvius dataset (train and test), F1-score for the fully labeled synthetic Irish and UK datasets. Instances shows the number of positive and negative/unlabeled samples. Clustering on PU–contrastive representation outperforms the histogram peak detection baseline to identify EV and non-EV households.

5 Experimental Results

In this section, we describe the datasets, training details, baselines, evaluation metrics, and report the results. The experiments aim to demonstrate that the proposed pipeline effectively learns discriminative representations of EV charging behavior from raw aggregated household consumption, even when only a limited number of positive labels are available and the majority of the data remains unlabeled.

Fluvius Dataset We use a private dataset provided by Fluvius, the Flemish DSO in Belgium. It contains smart meter electricity consumption data from households throughout 2023, recorded at 15-minute intervals. Due to the EV registration process (households self-report or register their EVs), a subset of households are labeled as EV users (positives), while the remainder are unlabeled. The dataset includes approximately $3k$ households with EV and $5k$ unlabeled households.

Validation Datasets For validation of the clustering performance, where trusted labels are required, we use publicly available datasets as well:

- **Irish dataset:** One-year electricity consumption (2010) for 1,919 households (Commission for Energy Regulation (CER) 2025), resampled to 15-minute intervals. No households originally had EVs, so synthetic EV charging patterns were added with randomized EV charging power (2.3–7.4 kW), duration (4–20 hours), start hour (18:00–23:00), and number of charging days (150–300).
- **UK dataset:** Two-month electricity consumption for 292 households, 99 with EVs (Energy Systems Catapult 2023), sampled at 30-minute intervals. Data was resampled to 15-minute intervals and extended with zeros to create one-year time series for model input consistency, which is compatible with the temporal masking augmentation applied during training.

Training Setup The model is trained on the Fluvius dataset using the PU-contrastive learning framework described in Section 4. We use a ResNet-18 backbone modified for two-dimensional heatmap inputs, followed by a fully connected projection head with 128 hidden dimension. The model is trained for 200 epochs with a batch-size of 256, using the Adam optimizer with an initial learning rate of 10^{-4}. The temperature τ in the puNCE loss is also tuned to 0.07, and the estimated fraction of positives in the unlabeled set π is set to 0.15 based on the knowledge of experts.

Evaluation Metrics For the Fluvius PU dataset, only recall is computed as the evaluation metric on the positive-labeled households, as the true labels for the unlabeled set are unknown. For the Irish and UK datasets, where fully labeled data is available (with synthetic EV patterns added in the Irish case), F1-score is used to evaluate the performance.

Evaluation Setup After training, we extract latent representations for all households in the validation datasets. We discard the projection head since it is specifically optimized for the contrastive loss, not for general representations and removing it often gives better representation for downstream task (Chen et al. 2020). The representations are normalized and reduced to two dimensions using t-SNE to mitigate the curse of dimensionality. We then perform clustering using k-means and spectral clustering with $k = 2$ as the number of clusters, assigning cluster labels based on the majority of known positives within each cluster. For spectral clustering, we tune the kernel variance γ to achieve the best performance. This downstream evaluation tests whether the PU-contrastive learned representations are truly discriminative for EV presence.

Baseline We compare our PU-contrastive approach with a histogram-based peak detection baseline, inspired by (Neubert et al. 2022), where histogram peaks in aggregated load data are used as informative features to detect potential EV charging events. Notably, the method currently used by Fluvius DSO to identify households with EV is based on this approach. Figure 3 illustrates this feature by comparing the kernel density estimates (KDE) of two exemplary households with and without registered EV. To implement this, we first compute a kernel density estimation (KDE) of the consumption and then detect local maxima within a manually defined range corresponding to expected EV charging power. A threshold is applied to distinguish significant peaks from background fluctuations.

Reporting Results Results are reported on separate train/test splits for Fluvius, and synthetic Irish and UK datasets which can be found in Table 1 and Figure 4. The t-SNE visualizations show that the learned representations clearly separate EV and non-EV households, confirming their discriminative power. Some overlap remains in the

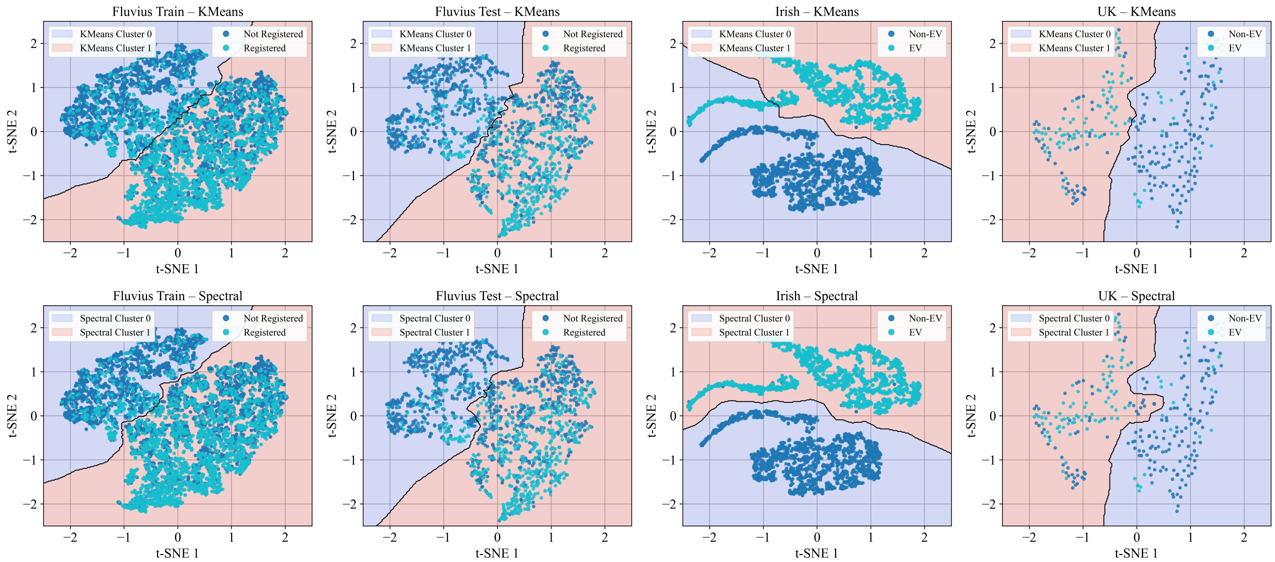

Figure 4: t-SNE visualization of learned representations with cluster boundaries. Top and bottom rows show k-means and spectral clustering on Fluvius, synthetic Irish, and UK datasets. Points are colored by EV registration status (Fluvius) or EV presence (Irish, UK), with cluster boundaries and predicted labels overlaid. The figure shows that both clustering methods roughly separate EV and non-EV households. In Fluvius, overlap occurs as registered EVs during observation year show weaker charging patterns, and some non-registered households may actually have EVs. The UK dataset shows mixed clusters due to the short two-month observation period, limiting the ability to distinguish EV charging patterns. The clustering of learned representations still outperforms baseline model in identifying potential EV households. The synthetic Irish dataset shows a clear separation between EV and non-EV households.

Fluvius dataset. This can be attributed to households that registered an EV only during the observation year, resulting in weaker charging patterns that place them closer to the non-EV cluster, while some households labeled as non-registered may actually own EVs which shifts them toward the EV cluster. In the UK dataset, the overlap is largely due to the limited two-month observation period, which restricts the visibility of consistent charging behavior. Despite these limitations, the overall clustering structure demonstrates that the proposed representations capture meaningful distinctions between EV and non-EV households and are more effective than baseline feature for distinguishing EVs.

6 Conclusion

We present a PU-contrastive learning approach for detecting electric vehicle (EV) adoption using real-world smart meter data. Our study leverages a private dataset obtained through our collaboration with Fluvius, the Flemish DSO, where only households with registered EV are labeled, and the majority of households remain unlabeled. This setup reflects practical deployment conditions, removing the need for costly or intrusive labeling of households. By learning latent representations from heatmaps of aggregated electricity consumption, our model effectively distinguishes EV users from non-EV households, even in the absence of negative labels. The resulting representation can be clustered to infer EV adoption. This framework offers DSOs a scalable, non-intrusive tool to identify households with EV, enabling them to anticipate localized increases in electricity demand caused by EVs and better manage grid resources. Such insights can improve load forecasting, guide infrastructure upgrades, and inform strategic planning. Given its practicality and strong performance, we are working toward deploying this approach within Fluvius to support data-driven decision-making for low-voltage grid operations.

Acknowledgments

This research is funded by: KU Leuven (Research Fund (projects IDN/24/013, iBOF/23/064), Industrial Research Fund and several LRD bilateral industrial projects); Flemish Government Agencies (FWO: TBM T001725N and 11K5623N; EWI: the Flanders AI Research Program VLAIO: CSBO (HBC.2021.0076)); FOD Volksgezondheid, Veiligheid van de Voedselketen en Leefmilieu: AI-943; EU: This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 885682). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or ERC. Neither the European Union nor the granting authority can be held responsible for them.

References

Acharya, A.; Sanghavi, S.; Jing, L.; Bhushanam, B.; Choudhary, D.; Rabbat, M.; and Dhillon, I. 2024. Positive Unlabeled Contrastive Learning.

Aly, H.; Al-Ali, A.; Al-Ali, A.; and Malluhi, Q. 2024. RES-EV: Identifying EV-Households under High AC Load Using a Residual-Based Model. In *2024 IEEE 8th Energy Conference (ENERGYCON)*, 1–6.

Becker, T.; Smet, R.; Macharis, B.; and Vanthournout, K. 2025. Effect of electric vehicles, heat pumps, and solar panels on low-voltage feeders: Evidence from smart meter profiles. *Sustainable Energy, Grids and Networks*, 42: 101705.

Bekker, J.; and Davis, J. 2020. Learning from positive and unlabeled data: a survey. *Machine Learning*, 109: 719–760.

Chen, T.; Kornblith, S.; Norouzi, M.; and Hinton, G. 2020. A Simple Framework for Contrastive Learning of Visual Representations. arXiv:2002.05709.

Commission for Energy Regulation (CER). 2025. CER Smart Metering Project - Electricity Customer Behaviour Trial, 2009-2010.

Criado-Ramón, D.; Ruiz, L. G. B.; Iruela, J. R. S.; and Pegalajar, M. C. 2024. A Novel Non-Intrusive Load Monitoring Algorithm for Unsupervised Disaggregation of Household Appliances. *Information*, 15(2).

Energy Systems Catapult. 2023. Spotting Low Carbon Technologies: Charlotte Avery. https://es.catapult.org.uk/insight/spotting-low-carbon-technologies-charlotte-avery/. Accessed: 2025-08-01.

Ghaffar, M.; Sheikh, S. R.; Naseer, N.; Din, Z. M. U.; Rehman, H. Z. U.; and Naved, M. 2022. Non-Intrusive Load Monitoring of Buildings Using Spectral Clustering. *Sensors*, 22(11).

Hoffmann, V.; and Fesche, B. I. 2019. AUTOMATED DETECTION OF ELECTRIC VEHICLES IN HOURLY SMART METER DATA. In *25th International Conference on Electricity Distribution*, 3–6.

Iwana, B. K.; and Uchida, S. 2021. An empirical survey of data augmentation for time series classification with neural networks.

Maaten, L. V. D.; and Hinton, G. 2008. Visualizing Data using t-SNE. Technical report.

Martin, C.; Ke, F.; and Wang, H. 2023. Non-Intrusive Load Monitoring for Feeder-Level EV Charging Detection: Sliding Window-based Approaches to Offline and Online Detection. arXiv:2312.01887.

Neubert, M.; Gnepper, O.; Mey, O.; and Schneider, A. 2022. Detection of Electric Vehicles and Photovoltaic Systems in Smart Meter Data. *Energies*, 15.

van den Oord, A.; Li, Y.; and Vinyals, O. 2019. Representation Learning with Contrastive Predictive Coding. arXiv:1807.03748.

Vavouris, A.; Garside, B.; Stankovic, L.; and Stankovic, V. 2022. Low-Frequency Non-Intrusive Load Monitoring of Electric Vehicles in Houses with Solar Generation: Generalisability and Transferability. *Energies*, 15(6).

Wang, X.; Tang, G.; Wang, Y.; Keshav, S.; and Zhang, Y. 2022. EVSense: a robust and scalable approach to non-intrusive EV charging detection. In *Proceedings of the Thirteenth ACM International Conference on Future Energy Systems*, e-Energy '22, 307–319. New York, NY, USA: Association for Computing Machinery. ISBN 9781450393973.

Yin, Y.; Xu, B.; Zhong, Y.; Bao, T.; and Wang, P. 2023. A Non-Intrusive Load Monitoring Model for Electric Vehicles Based on Multi-Kernel Conventional Neural Network. *World Electric Vehicle Journal*, 14.

Zhang, Z.; Son, J. H.; Li, Y.; Trayer, M.; Pi, Z.; Hwang, D. Y.; and Moon, J. K. 2014. Training-free non-intrusive load monitoring of electric vehicle charging with low sampling rate. In *IECON 2014 - 40th Annual Conference of the IEEE Industrial Electronics Society*, 5419–5425.