ChatGPT

Liquid Neural Networks and Continuous-Time Models in Time-Series Tasks

Liquid Neural Networks (LNNs) are a class of continuous-time recurrent models that adapt their dynamics based on input data, unlike conventional RNNs with fixed parameters [1] [2]. They originated from **liquid time-constant networks (LTCs)**, which use neurons governed by ordinary differential equations (ODEs) with input-dependent time constants [3] [4]. This gives LNNs a "liquid" behavior: the neuron's response (time constant) varies with inputs, enabling the network to continually adjust to new data. Recent advances like **Closed-form Continuous-time (CfC) networks** eliminate the need for numerical ODE solvers by finding an analytical solution for the dynamics [5] [6]. These developments improve computational efficiency while preserving the expressiveness of continuous-time models. Below, we survey impactful recent research applying LNNs and related ODE-based models in **ECG classification**, **human activity recognition (HAR)**, and **speech recognition**, then highlight open research gaps and opportunities.

LNNs in ECG Classification

Benchmark Datasets: A widely-used benchmark for ECG arrhythmia classification is the **China Physiological Signal Challenge 2018 (CPSC-2018)** 12-lead ECG dataset [7], designed for automated arrhythmia detection. Another large dataset is the **Telehealth Network of Minas Gerais (TNMG)**, containing over 2.3 million 12-lead ECG samples labeled with rhythm abnormalities [8]. Recent top-tier works often evaluate on CPSC for testing generalization, sometimes using TNMG or other collections for training [9] [10]. Other common ECG benchmarks include legacy datasets like **MIT-BIH Arrhythmia** and newer ones like **PTB-XL**, but recent LNN studies focus on multi-lead clinical datasets (e.g. CPSC, Chapman University's dataset, etc.) to address multi-class arrhythmia detection.

Recent LNN-Based Models: Huang *et al.* (2024) introduced two LNN-infused models for multi-class arrhythmia detection on resource-limited hardware [11] [12]. They proposed:

- **ConvLSTM2D-LTC (CLTC):** a hybrid model combining 2D convolutional layers, a ConvLSTM layer, and an LTC RNN cell.
- **ConvLSTM2D-CfC (CCfC):** a similar architecture replacing the LTC cell with a Closed-form Continuous-time RNN cell.

Both models were trained on a balanced subset of the TNMG dataset and evaluated on the held-out TNMG data and the CPSC-2018 test set [7] [8]. The table below summarizes their performance:

Model (Huang et al. 2024)	Dataset	Accuracy / F1	AUROC	Venue (Year)	Notes
ConvLSTM2D-**LTC** (CLTC)	TNMG (internal)	F1 ≈ 0.827	0.961	J. Cardiovasc. Transl. Res. (2024) [13] [13]	Multi-class arrhythmia; Deployed on STM32 MCU
ConvLSTM2D-**CfC** (CCfC)	TNMG (internal)	F1 ≈ 0.828	0.963	J. Cardiovasc. Transl. Res. (2024) [13] [13]	Slightly higher accuracy; ~40% faster training [14]
Generalization (CCfC)	CPSC-2018 (external)	F1 ≈ 0.72	0.91	–	Cross-dataset evaluation [15] (multi-class)

Table 1: Performance of LNN-based models on ECG classification benchmarks. Both models achieved similar within-distribution performance (~0.82 F1 on TNMG) [13]. Notably, when tested on the external CPSC dataset, the LNN models still attained F1 ~0.72 [15], demonstrating good generalization across different patient populations. The **AUROC** remained high (>0.90), indicating strong discriminative ability even on unseen data [15]. These models were also successfully deployed on a microcontroller (using only ~240KB RAM and ~96KB flash) – a testament to LNNs' compactness and efficiency [16] [17].

Novelty: This work is impactful as one of the first to apply LNNs (LTC/CfC) to clinical ECG tasks. The models tolerate noisy or missing leads better than conventional deep networks: e.g. the LTC-based model handled dropped ECG channels more robustly, and the CfC-based model maintained slightly higher overall accuracy [18]. The adaptive continuous-time neurons likely confer resilience to signal perturbations. Moreover, prior tiny-ECG models on edge devices were mostly limited to single-rhythm (binary) classification. For example, a NAS-automated tiny model (SenSys 2022) reached F1 ≈ 0.83 for atrial fibrillation detection on-device [19], and an ultra-light MLP co-inference model achieved F1 > 0.80 for AF on a wearable chip [20]. Huang *et al.* (2024) surpass these by tackling **multi-class** arrhythmia classification with one model [21], extending edge-AI ECG analysis beyond binary detection.

Venue: The work by Huang *et al.* appeared in *Journal of Cardiovascular Translational Research* (2024) [22] as an open-access article, reflecting its interdisciplinary appeal (deep learning and cardiology). Another related paper by the same group proposed **S4D-ECG**, a shallow state-space model for arrhythmias, which obtained F1 ≈ 0.81 using a simplified S4 sequence model [23]. While not an LNN, S4D-ECG underscores the interest in state-space and ODE-inspired models for ECG. Overall, LNN-based approaches are just emerging in top venues; there is room to push them into AI and medical conferences (e.g. AAAI, NeurIPS healthcare workshops) with further performance gains and validation.

LNNs in Human Activity Recognition (HAR)

Benchmark Datasets: The canonical dataset for HAR is the **UCI Human Activity Recognition (HAR) using Smartphones** dataset [24]. It contains 6,554 time series of smartphone inertial sensor readings from 30 subjects performing daily activities (walking, sitting, lying, etc.) [25]. Each sequence is a multivariate time series of 561 features per time-step (accelerometer and gyroscope signals) with a label for the activity being performed [26]. Many studies use the UCI HAR dataset due to its public availability and inclusion in early time-series ML benchmarks. Other HAR datasets (e.g. **Opportunity**, **PAMAP2**) are used in activity recognition research, but top-tier continuous-time modeling papers have focused on UCI HAR as a representative benchmark for irregularly-sampled sequences [27].

Recent Results: The most notable application of LNNs to HAR is by Hasani *et al.* (2022) in their **Closed-form Continuous-time Neural Networks** paper (Nature Machine Intelligence) [28] [29]. They evaluate various CfC model variants on the UCI HAR dataset and compare against numerous sequential models (both discrete-time RNNs and neural ODE baselines). The task is framed as *per time-step classification* – each time sample in the sequence is labeled with an activity, following the preprocessing of Rubanova *et al.* (2019) to simulate irregular sampling [27]. Key results from their **Table 2** are:

Model	HAR Accuracy (% per timestep)	Training Time/Epoch
Latent ODE (ODE solver) [30]	84.6 ± 1.3	8.49 min
ODE-RNN [31]	82.9 ± 1.6	3.15 min
GRU-D (ICLR 2018) [32]	80.6 ± 0.7	0.15 min
CfC (closed-form RNN) [33]	84.9 ± 0.4	**0.084 min**
CfC-noGate [34]	85.6 ± 0.3	0.093 min
CfC-mmRNN (multi-memory) [35]	86.0 ± 0.3	0.128 min
CfC-S (single-step) [36]	**87.0 ± 0.5**	0.097 min

Table 2: HAR classification accuracy using LNNs vs. baselines [37] [36]. Bold indicates the best accuracy. All CfC variants outperform the classical and ODE-based RNN baselines on this benchmark. Notably, the best LNN (CfC-S) reaches **87.0% accuracy**, exceeding the previous state-of-the-art Latent ODE's 84.6% [30] [36]. This is a sizable improvement on this task. Moreover, the closed-form models train **orders of magnitude faster** – CfC-S trains in ~0.1 minutes/epoch, over **8750% faster** than the Latent ODE which requires solving stiff ODEs at each step [38] [39]. This dramatic speedup is because CfC analytically integrates the dynamics, avoiding the overhead of numerical solvers [40].

Novelty: The CfC models demonstrate that LNNs can excel in HAR by handling irregular timings and long sequences efficiently. They achieve top accuracy *and* faster computation by virtue of a closed-form solution that sidesteps the numerical integration bottleneck [38]. The CfC variants also incorporated innovations like gating (CfC-noGate removes gates to test their effect) and a multi-timescale extension (CfC-mmRNN introduces multiple recurrent units for different frequency components) [38]. These achieved competitive accuracy, suggesting LNNs can be enhanced with gating or multi-memory structures to further boost performance. The **novelty** lies in proving continuous-time models can match or beat discrete models on a standard HAR task, and in quantifying the huge efficiency gain [39] [41].

Venue: This work was published in *Nature Machine Intelligence* (Nov 2022) [28], a top-tier venue, highlighting its impact. Additionally, earlier continuous-time models like **Latent ODE** (Rubanova *et al.*, NeurIPS 2019) had also tested on HAR, but with lower accuracy (around 80–85% as seen above) due to reliance on ODE solvers [31] [30]. The CfC approach closed that gap.

Overall, UCI HAR is a relatively solved dataset in traditional deep learning (many CNN/LSTM models exceed 95% accuracy on segment-level classification). The continuous-time setting with per-timestep labels is more challenging, explaining the ~87% ceiling here. A research opportunity is to apply LNNs to *richer HAR datasets* (e.g. wearable multi-sensor data from daily life, or HAR with irregular sampling from devices) where their adaptability to time-varying inputs could shine. So far, LNNs have proven effective on the benchmark, establishing a new state-of-the-art for continuous-time HAR modeling [38].

LNNs in Speech Recognition

Benchmark Datasets: In speech tasks, continuous-time models have been explored mainly on *keyword spotting* and small vocabulary datasets. A popular benchmark is the **Google Speech Commands** dataset, which includes short (one-second) audio clips of spoken words (35 classes in the “full” version) for classification. This task tests sequence models’ ability to process raw audio waveforms or spectrograms. Large-vocabulary speech recognition (e.g. LibriSpeech) has not yet seen extensive use of LNNs in literature, likely because Transformer architectures currently dominate there. Instead, researchers have evaluated LNNs on Speech Commands or toy speech datasets to validate their sequence learning capability.

Recent Results: A cutting-edge example is **Liquid-S4**, proposed by Hasani *et al.* and presented at ICLR 2023 [42] [43]. Liquid-S4 combines the *structured state-space model* (SSM) approach of S4 (Gu *et al.*, NeurIPS 2022) with liquid time-constant continuous-time dynamics. In essence, it replaces S4’s fixed linear state matrix with an **LTC-based state transition** that adapts to inputs, while retaining the high-order polynomial projection and convolutional kernel of S4 [44] [45]. This hybrid achieves state-of-the-art results across long-range sequence benchmarks. On the **Speech Commands (35-word)** task, Liquid-S4 reached **96.78% accuracy** on the test set [43]. This slightly outperforms the original S4 model, while using **30% fewer parameters** [43]. In other words, by injecting “liquid” neurons into the SSM, the model became more compact and generalizable. Liquid-S4 also outperformed various RNNs, CNNs, and Transformers on this speech benchmark [46] [47]. For context, S4 had already reported ~95-96% accuracy on Speech Commands full dataset [48] [49], which was a record among sequence models; Liquid-S4 edged this out to set a new state-of-art.

Another relevant study, Wang *et al.* (2024), introduced an **Attention-Conv-CfC (AC-CfC)** model for speech emotion recognition (on EEG signals) [50]. While not speech *recognition*, it combined 1D CNNs and CfC units with an attention mechanism to capture temporal dynamics in multi-channel signals. This suggests LNNs are starting to appear in audio-related domains, augmented with attention for improved performance. However, applications of LNNs to full automatic speech recognition (ASR) (continuous transcription tasks) are still *underexplored*. No major ASR benchmark results with LNNs have been reported in top conferences to date, to our knowledge. It remains an open area whether LNNs can replace or augment Transformers for large-scale speech recognition, where sequence lengths are very long and real-time adaptability could help with varying speech rates or noise conditions.

Novelty: The Liquid-S4 result is impactful as it merges continuous-time adaptability with SSM’s long-range memory. Achieving ~96.8% on Speech Commands with a smaller model demonstrates **efficiency and robustness** [43]. The LNN component allows the model to dynamically adjust to different pronunciations or audio conditions, potentially improving noise robustness (an aspect to be quantified in future work). The novelty also lies in bridging two advanced sequence modeling paradigms (LTC RNNs and S4). This paves the way for **LNN-Transformer or LNN-SSM hybrids** that could scale to tasks like speech and language. The research was presented at ICLR 2023 (poster) [42], underscoring its relevance to the machine learning research community.

Research Gaps and Opportunities

Despite encouraging progress, several gaps remain where **new LNN-based architectures** could contribute novel insights and improved performance:

- **Multi-Domain Benchmark Expansion:** Thus far, LNNs have shown promise on relatively controlled benchmarks (arrhythmia classification, UCI HAR, Speech Commands). There is room to apply them to *broader and more complex datasets* that are actively benchmarked in top conferences. For instance, in **ECG analysis**, LNNs could be tested on multi-label diagnosis tasks (detecting multiple arrhythmias per record) using datasets like PTB-XL or the PhysioNet Challenge 2021 data. These tasks often use transformer or CNN models; an LNN's compactness and continuous adaptability might yield better efficiency or personalization. In **HAR**, datasets like Opportunity or newer multimodal datasets (e.g. combining wearable sensors and video) present an opportunity – LNNs can inherently handle irregular or missing sensor readings, a common issue in real-world HAR, which many Transformer models struggle with. Extending LNN evaluations to such benchmarks could validate their advantages in handling **uneven sampling and multi-rate data**.

- **Outperforming RNNs/Transformers in Underexplored Tasks:** LNNs excel in scenarios with **continuous input streams or dynamically changing conditions** [1] [51]. One underexplored area is **online learning and concept drift** in time-series. Because LNN neurons update their time-constant based on input, they might adapt more gracefully to changes in data statistics over time (e.g. a person's heart rhythm shifting due to medication, or a user performing new activities with a wearable). Research could focus on tasks like *early anomaly detection* or *lifelong activity recognition*, where models must adapt to new classes or drifts – LNNs could maintain stability while learning on the fly. Another opportunity is **robotics and control**, which, though beyond the scope of this query, have seen initial success with LNNs (e.g. Neural Circuit Policies for autonomous driving [52]). Incorporating LNNs into reinforcement learning for control or into **CVPR-style video action recognition** (treating video frames as continuous-time signals) could yield models that handle variable frame rates and motion dynamics better than discrete counterparts.

- **Efficiency on Edge and Low-Power Devices:** A clear advantage of LNNs is their parsimonious nature – fewer neurons can achieve rich dynamics [53] [54]. This suits them for embedded applications (wearables, IoT sensors) where memory and power are limited. The current research shows feasibility (e.g. ~240 KB RAM for an ECG LNN model) [16] [17], but **optimizing LNN inference** is an open problem. Future contributions could explore quantization or compiler optimizations for LNN ODEs, enabling even faster and more energy-efficient deployment. Moreover, **scalability** remains partly unexplored: LNNs on microcontrollers work well, but can we scale up an LTC/CfC network to large GPU clusters for big data? The *Liquid-S4* work addressed sequence length scaling, but scaling network *depth/width* without losing stability is an open question. Research can target combining LNN cells with deep architectures – for example, using LNN layers within a deep CNN or Transformer to inject continuous-time reasoning, or stacking LNN layers for hierarchical temporal processing. Ensuring training remains stable as we deepen LNN architectures will be crucial (perhaps borrowing techniques from Neural ODE stability analysis or using spectral normalization on LNN gates).

- **Adaptability and Robustness:** Many claims about LNNs highlight their robustness to noisy or perturbed inputs [51] [16]. Empirical studies should rigorously test this. For instance, adding extreme noise or simulating sensor failures in HAR, or adversarial perturbations in ECG, would reveal if LNNs truly maintain performance better than LSTMs/Transformers. Initial evidence is positive (the NCP-based ECG model handled transient disturbances better than an LSTM [51]), but quantifiable benchmarks for robustness are needed. This could open publishable insights in venues like ICLR (which in recent years have had a focus on robust and **adaptive deep learning**). Additionally, **personalization** – adapting a trained model to a new user or patient with minimal data – could leverage LNN adaptability. One could envision an LNN that, after deployment, *continues to fine-tune its time-constants* in response to a user's data (a kind of continuous transfer learning). This is relatively unexplored and could be a unique selling point for LNNs in health and wearable computing research.

- **Integrating Memory and Attention:** LNNs in current form are excellent at local, continuous dynamics, but less so at capturing very long-term dependencies unless augmented (e.g. CfC-mmRNN with multiple recurrent units [38], or Liquid-S4 using SSM kernels [55]). Research could explore *hybrid architectures* where LNNs are combined with attention mechanisms or external memory. For example, an **LNN-Transformer hybrid** might use a CfC or LTC layer to encode input streams into a hidden state that is fed into a self-attention block for capturing global context. This way, one might get the best of both worlds: LNN's continuous adaptability and a Transformer's long-range sequence modeling. Such ideas align with trends in top conferences (e.g. **continuous-time transformers** [56] have been proposed for irregular time-series). An *attention-based CfC* was already used for EEG emotion recognition, indicating feasibility [50]. Extending this to, say, speech recognition (where an LNN could preprocess audio into more stable features for a Transformer to decode) is an exciting direction.

In summary, **Liquid Neural Networks are a nascent yet powerful paradigm** for time-series modeling. They have proven their merit on key benchmarks in ECG classification, HAR, and speech commands, often achieving state-of-the-art results with fewer parameters and greater robustness (e.g. 96.8% accuracy on Speech Commands [43], 87% accuracy on HAR [36], and efficient ECG anomaly detectors on-device [57] [58]). The research gaps highlight that many application domains and integration schemes remain unexplored. By targeting datasets with open challenges (e.g. multi-label ECG, real-world HAR, noisy speech) and focusing on LNN's strengths (efficiency, adaptability, continuous dynamics), researchers can carve out **viable, novel contributions**. For instance, a paper proposing a **multi-scale LNN** (drawing inspiration from biology or the recent "Memory Enhanced LTC" concept) that outperforms Transformers on an irregular physiological time-series dataset could be impactful. Likewise, demonstrating an LNN-based model significantly reducing computation or improving robustness in a critical task (like wearable seizure detection or personalized speech interfaces) would attract top-tier attention.

Most Promising Directions: Based on our survey, promising research directions include:

- *Applying LNNs to new benchmarks:* e.g. integrating **Liquid-S4** models for long-term ECG monitoring or multivariate clinical time-series to set new benchmarks in health analytics (building on its success in sequence modeling [46] [47]).
- *Efficiency and interpretability in edge scenarios:* devising **tiny LNN architectures** for wearable sensors that maintain accuracy with ultra-low power. The inherent sparsity and simplicity of some LNN

variants (like Neural Circuit Policies) already aids interpretability [59] [60] – future work can emphasize explainable decisions (crucial for healthcare).

- *Hybrid models:* combining **continuous-time LNN cells with discrete-time architectures** (Transformers or CNNs) to handle both fine-grained temporal dynamics and high-level sequence structure. This could address tasks like activity segmentation or event forecasting where both short-term and long-term patterns matter.
- *Theoretical understanding and new variants:* investigating the theoretical properties of LNNs (stability, expressive power) to guide architecture design. For example, recent work introduced **Liquid Structural State-Space Models** that linearize LTCs and achieved SOTA on long-range tasks [45] [46] – further simplifying LNN equations (e.g. finding exact solution forms for more complex neurons) could remove remaining bottlenecks. Additionally, introducing new components (attractor dynamics, as in a 2025 preprint with memory neurons) can help LNNs capture multiple timescales [61] [62].

By addressing these gaps, researchers can push LNNs into the mainstream of sequence learning, potentially achieving breakthroughs such as *real-time adaptive speech recognition systems, robust health monitors that learn from patients over time, and energy-efficient sensor networks with on-device intelligence*. Each of these would represent a publishable advancement in venues like NeurIPS, ICLR, or CVPR, especially as the community seeks models that are **smaller, more adaptable, and closer to physical dynamics** than today's static deep networks [53] [54].

Sources: Recent key papers and results are cited throughout, including Hasani *et al.* (2021, 2022) [63] [38], Huang *et al.* (2024) [11] [13], and others integrating LNNs into various domains. These illustrate the state-of-the-art performance and point toward how *liquid* neural networks can reshape time-series AI.

---

[1] [52] [53] [54] "Liquid" machine-learning system adapts to changing conditions | MIT News | Massachusetts Institute of Technology
https://news.mit.edu/2021/machine-learning-adapts-0128

[2] [3] [4] [2006.04439] Liquid Time-constant Networks
https://ar5iv.org/pdf/2006.04439.pdf

[5] [6] [7] [8] [9] [10] [14] [16] [17] [51] [59] [60] Efficient Edge-AI Models for Robust ECG Abnormality Detection on Resource-Constrained Hardware | medRxiv
https://www.medrxiv.org/content/10.1101/2023.08.31.23294925v1.full-text

[11] [12] [13] [15] [18] [19] [20] [21] [22] [23] [57] [58] Efficient Edge-AI Models for Robust ECG Abnormality Detection on Resource-Constrained Hardware | Journal of Cardiovascular Translational Research
https://link.springer.com/article/10.1007/s12265-024-10504-y

[24] [25] [26] [27] [28] [29] [38] [39] [40] [41] Closed-form continuous-time neural networks | Nature Machine Intelligence
https://www.nature.com/articles/s42256-022-00556-7

[30] [31] [32] [33] [34] [35] [36] [37] [63] Table 2 Human activity recognition, per time-step classification
https://www.nature.com/articles/s42256-022-00556-7/tables/2

[42] [43] Liquid Structural State-Space Models | OpenReview
https://openreview.net/forum?id=g4OTKRKfS7R

44 45 46 47 48 49 55 arxiv.org
https://arxiv.org/pdf/2209.12951

50 AC-CfC: An attention-based convolutional closed-form continuous ...
https://www.researchgate.net/publication/382772549_AC-CfC_An_attention-based_convolutional_closed-form_continuous-time_neural_network_for_raw_multi-channel_EEG-based_emotion_recognition

56 Continuous-Time Transformer for Irregular Time Series Modeling
https://nips.cc/virtual/2023/poster/71304

61 62 Time Series Analysis Using Memory Enhanced Liquid Neural Network by Srinivas M, Abhijeet Kumar, RBV Subramaanyam :: SSRN
https://papers.ssrn.com/sol3/Delivery.cfm/2cc4ef1c-1419-4c90-9cd4-5f77915ed838-MECA.pdf?abstractid=5253744&mirid=1