

Berdasarkan sumber-sumber yang tersedia, berikut adalah kurasi referensi yang paling relevan untuk **data augmentation** (khususnya WGAN dan variannya) dalam konteks **Network Traffic, IoT, dan Reinforcement Learning (RL)**.

Saya telah mengelompokkannya berdasarkan kontribusi spesifik untuk paper Anda agar mudah disitusi:

## 1. Referensi Utama untuk Augmentasi Data Trafik (WGAN-GP & LLM)

Ini adalah referensi paling "game-changer" untuk masalah dataset Anda yang statis.

- **Alyoubi et al. (2025) - "Network Traffic Data Augmentation Using WGAN Model Guided by LLM"** 1:
  - **Relevansi:** Ini adalah referensi terbaik untuk solusi Anda. Paper ini mengatasi masalah dataset trafik IoT yang tidak seimbang dengan menggunakan **WGAN** yang dipandu oleh **Large Language Model (LLM)** dan **Graph-Conditioned**.
  - **Ide untuk Paper Anda:** Anda bisa mengadopsi metode mereka dalam menggunakan *Pearson correlation* dan *Mutual Information* untuk membangun graf hubungan antar fitur (misal: hubungan antara rx\_mbps dan rx\_pps di dataset Anda) agar data buatan tetap mematuhi aturan protokol jaringan, bukan hanya sekadar angka acak 2, 3.
  - **Metode:** Menggabungkan statistik struktural (Graph) dengan aturan semantik (LLM) untuk memaksa generator WGAN menghasilkan trafik yang valid secara protokol 2.
- **Bouzeraib et al. (2025) - "Enhancing IoT Intrusion Detection Systems Through Horizontal Federated Learning and Optimized WGAN-GP"** 4:
  - **Relevansi:** Sangat relevan karena menggunakan **WGAN-GP (Gradient Penalty)** secara spesifik untuk IoT Network Intrusion Detection Systems (NIDS). Ini cocok dengan skenario *testbed* Anda yang melibatkan keamanan/anomali pada slice jaringan.
  - **Poin Kunci:** Menunjukkan bahwa WGAN-GP lebih stabil daripada GAN biasa dalam menangani *mode collapse* (masalah di mana generator hanya menghasilkan satu jenis output) 5, 6.
  - **Riaz et al. (2025) - "A novel ensemble Wasserstein GAN framework..."** 7:
    - **Relevansi:** Memperkenalkan **EO-WGAN (Enhanced Optimization WGAN)**. Mereka menggunakan pendekatan dua tahap: pertama menggunakan SMOTE untuk interpolasi kasar, lalu WGAN untuk memperhalus distribusi data agar realistik untuk lingkungan *Industrial IoT* (IIoT) 7, 8.
    - **Ide:** Jika WGAN murni terlalu berat, pendekatan *hybrid* SMOTE + WGAN ini bisa menjadi alternatif metodologi untuk paper Anda.

## 2. Referensi untuk Time-Series Traffic (Penting untuk RL)

Karena RL membutuhkan data yang berurutan (*state \$t, t+1, \dots\$*), augmentasi data tabular biasa seringkali merusak pola waktu.

- **König et al. (2023) - "A LSTM-GAN Algorithm for Synthetic Data Generation of Time Series..."** 9:
  - **Relevansi:** Mengusulkan penggabungan **LSTM (Long Short-Term Memory)** dengan **WGAN**.
  - **Ide:** Karena dataset Anda memiliki timestamp dan RL membutuhkan urutan waktu (sequential decision making), Anda tidak bisa hanya mengacak baris. Referensi ini

memvalidasi penggunaan arsitektur LSTM-WGAN untuk menghasilkan data *time-series* yang koheren, yang sangat krusial agar agen RL Anda bisa belajar pola trafik (misal: *burst* trafik kamera di Port 2) 9.

### 3. Referensi Survei & Teori Pendukung (Untuk Bab Related Work)

Gunakan ini untuk membuktikan bahwa pendekatan Anda valid dan *state-of-the-art*.

- **Survey on GANs for IDS (2025) 10:**
- **Relevansi:** Ini adalah survei komprehensif tentang penggunaan GAN dalam keamanan jaringan dan IoT.
- **Poin Kunci:** Membahas varian **Conditional GAN (CGAN)** dan **Wasserstein GAN** untuk mengatasi *class imbalance* pada trafik jaringan 10, 11. Anda bisa mengutip ini untuk menjustifikasi mengapa Anda memilih WGAN daripada teknik augmentasi lama.
- **Iftikhar & Ali (2025) - "Investigating the Impact of Hyper Parameters..." 12:**
- **Relevansi:** Studi teknis tentang pengaruh *learning rate*, *epochs*, dan persentase augmentasi pada performa WGAN-GP. Sangat berguna saat Anda menulis bagian *Experimental Setup* atau *Hyperparameter Tuning* di paper Anda 12.

### 4. Referensi Penghubung Augmentasi dengan RL (Decision Making)

Bagaimana data buatan ini membantu Agen RL (DQN/PPO)?

- **Xia et al. (2022) - "Genet: Automatic Curriculum Generation for Learning Adaptation in Networking" 13:**
- **Relevansi:** Meskipun tidak murni WGAN, paper ini membahas **Curriculum Learning** untuk RL di jaringan.
- **Ide Kontribusi:** Anda bisa menggunakan konsep ini untuk menjelaskan bahwa data augmentasi WGAN Anda digunakan untuk membuat "kurikulum" pelatihan bagi agen RL. Misalnya, latih agen PPO Anda dengan data normal dulu, lalu secara bertahap masukkan data sintetis "ekstrem" (macet/serangan) hasil WGAN agar agen menjadi lebih *robust* 14, 15. Ini sangat relevan dengan tujuan Anda membandingkan PPO dan DQN.

Ringkasan Rekomendasi untuk Paper Anda:

Untuk bagian **Data Augmentation & Preprocessing** di paper Anda, saya sarankan alur referensi berikut:

1. Gunakan **4 (Bouzeraib et al.)** dan **7 (Riaz et al.)** sebagai dasar utama penggunaan **WGAN-GP** untuk mengatasi ketidakseimbangan data pada trafik IoT.
2. Gunakan **1 (Alyoubi et al.)** sebagai referensi metode mutakhir untuk memandu WGAN menggunakan **aturan semantik (Graph/LLM)** agar data sintetis Port 4 (Healthcare) dan Port 2 (Kamera) tetap masuk akal secara protokol.
3. Gunakan **9 (König et al.)** untuk menjustifikasi penanganan fitur *time-series* (timestamp) agar agen RL bisa mempelajari pola temporal.
4. Hubungkan hasilnya ke RL dengan mengutip **13 (Xia et al.)**, menyatakan bahwa data augmentasi ini berfungsi sebagai lingkungan pelatihan yang diperkaya (*enhanced training environment*) untuk agen PPO/DQN Anda.