

Drone Attitude Estimation under Out-of-Distribution Sensor Noise: A Comparative Study of EKF and LSTM

Bu-Yeon Kim*, Seok-Won Lee**

*Student, Graduate School of Information, Communication and Technology, Ajou University, Suwon, Korea

**Professor, Dept. of Software and Computer Engineering, Ajou University, Suwon, Korea

[Abstract]

The Extended Kalman Filter (EKF) is widely used for drone attitude estimation, but its accuracy can degrade when sensor noise deviates from the assumed noise model. Long Short-Term Memory (LSTM) networks have been applied to IMU-based attitude estimation, yet out-of-distribution (OOD) generalization remains underexplored. This study examines whether an LSTM trained on low-noise data can achieve lower root mean square error (RMSE) than the EKF under unseen high-noise conditions. In the ArduPilot SITL, sensor noise was configured at four levels (L0-L3); a quaternion-output LSTM was trained on L0-L1 and evaluated on L2-L3. The LSTM achieved 21.4% lower mean RMSE than the EKF at L2 and 14.1% lower at L3. However, per-window MAX results indicate that this advantage did not extend to all error metrics. These findings suggest that the LSTM model is better viewed as a complement to the EKF rather than as its replacement under high-noise conditions.

▶ **Key words:** Drone attitude estimation, Sensor fusion, Extended Kalman filter, LSTM, Out-of-distribution generalization

[요 약]

확장 칼만 필터(EKF)는 드론 자세추정에 널리 사용되어 왔으나, 실제 노이즈가 사전에 가정한 모델 범위를 벗어나면 추정 정확도가 저하될 수 있다. IMU 기반 자세추정에 LSTM을 적용한 연구가 보고되었으나, 대부분 학습 노이즈 분포 안에서 평가되어 분포 외(Out-of-Distribution, OOD) 일반화는 충분히 검증되지 않았다. 본 연구는 저노이즈 구간으로만 학습한 LSTM이 학습에 사용되지 않은 고노이즈 구간에서도 EKF보다 낮은 오차를 유지할 수 있는지를 검증한다. ArduPilot SITL의 노이즈 파라미터를 조합하여 노이즈 강도를 4단계(L0-L3)로 분할하고, 저노이즈 구간(L0, L1)으로만 학습한 쿼터니언 출력 LSTM을 고노이즈 구간(L2, L3)에서 평가하였다. 평가 결과 LSTM의 평균 RMSE는 EKF 대비 L2에서 21.4%, L3에서 14.1% 낮았으며 윈도우 단위 최대 오차에서는 EKF가 더 안정적이었다. 이는 LSTM과 EKF가 평균 오차와 최대 오차에서 서로 다른 강점을 가져 보완적으로 활용될 수 있음을 보인 결과이다.

▶ **주제어:** 드론 자세추정, 센서 융합, 확장 칼만 필터, LSTM, 분포 외 일반화

• First Author: Bu-Yeon Kim, Corresponding Author: Seok-Won Lee

*Bu-Yeon Kim (bykim523@ajou.ac.kr), Graduate School of Information, Communication and Technology, Ajou University

**Seok-Won Lee (leesw@ajou.ac.kr), Dept. of Software and Computer Engineering, Ajou University

• Received: 2026. 05. 06, Revised: 2026. 05. 21, Accepted: 2026. 06. 17.

I. Introduction

재난 구조나 산업 점검을 비롯한 여러 분야에서 무인 항공기(Unmanned Aerial Vehicle, UAV)의 활용이 점차 늘고 있다[1, 2]. 자율비행 수요가 확대됨에 따라 핵심과제로 떠오른 것이 UAV 자세의 정확한 추정이며, 자세 표현에는 보통 롤·피치·요나 쿼터니언이 쓰인다[3, 4]. 비행 제어기가 안정적으로 이루어지려면 자세추정이 정확해야 한다. 오차가 쌓이면 의도한 자세에서 벗어나기 시작하고, 심한 경우 호버링 불안정이나 추락 위험으로 이어질 수 있으므로 자세추정 알고리즘은 평균 오차뿐 아니라 짧은 시간에 나타나는 큰 오차까지 함께 줄일 필요가 있다.

자세추정의 결과는 드론의 위치 제어와 경로 추종에도 영향을 준다. Fig. 1은 자세추정 오차가 비행 제어와 위치 안정성에 영향을 주는 흐름을 보여준다. 자세 오차는 추력 방향의 어긋남으로 이어지며, 호버링 상황에서는 위치 밀림으로, 경로 추종 상황에서는 목표 궤적 이탈로 나타날 수 있다. 그런데 자세추정 성능은 센서 노이즈와 바이어스, 외부 교란의 특성에 좌우된다. 실제 운용 환경에서는 센서 입력의 분포가 알고리즘 설계 시 가정된 범위를 벗어나는 일이 적지 않다. 저가형 MEMS IMU 사용, 주변 자기 교란, 기체 진동, 온도 변화에 따른 자이로 바이어스 누적 등이 대표적이다. 이러한 분포 외(Out-of-Distribution, OOD) 조건에서도 추정 오차가 급격히 커지지 않아야 실제 비행 제어에 활용할 수 있으며, 이를 위해서는 평균 성능뿐 아니라 OOD 조건에서의 강건성 검토가 필요하다.

자세추정에는 보통 9축 IMU(가속도계 3축, 자이로스코프 3축, 지자기센서 3축)가 사용되며[4], 다중 센서 융합 기법으로는 확장 칼만 필터(Extended Kalman Filter, EKF)가 오랫동안 자리잡아 왔다. EKF는 모델 기반 예측과 센서 측정값 보정을 결합한다는 장점이 있다. 다만 노이즈 공분산 튜닝에 민감하고, 외부 가속이나 자기교란을 자세 성분과 분리하기 어렵다는 문제가 반복적으로 지적되어 왔다[5, 6].

EKF가 요구하는 노이즈 모델 가정에 대한 의존도를 낮춘 대안으로 데이터 기반 접근이 주목받고 있다. IMU 신호 기반 자세추정에 적합한 구조로 알려진 것이 장단기 기억 신경망(Long Short-Term Memory, LSTM)이다. 시계열의 장기 의존성을 학습할 수 있는 구조이기 때문이다[7, 8]. 가속 조건에서 LSTM이 칼만필터·상보필터보다 낮은 추정 오차를 달성한 사례[5], 자기교란 환경에서 기존 필터 대비 강건한 결과를 낸 순환신경망 연구[6] 등이 그 가능성을 뒷받침한다. 다만 기존 연구들은 학습에 사용한

것과 비슷한 노이즈 조건에서 모델을 평가하는 데 그쳤고, 학습 범위를 넘어서는 OOD 노이즈 조건에서의 검증은 충분히 다루어지지 못하였다. 다중 센서 융합에 관한 종합적 고찰에서도 학습 기반 접근의 일반화 검증 필요성이 지적된 바 있다[9].

이 공백을 메우고자 OOD 평가 프레임워크를 제안한다. 노이즈 레벨은 ArduPilot SITL 환경에서 가속도계·자이로스코프·지자기센서의 노이즈 강도와 자이로 바이어스, 드리프트 속도를 조합하여 L0(노이즈 없음)부터 L3(극한 노이즈)까지 4단계로 정의하였다. LSTM 모델은 저노이즈 구간(L0, L1)으로만 학습한 뒤, 학습에 쓰이지 않은 고노이즈 구간(L2, L3)에서 EKF와 비교한다. 평가 절차에는 시드 10개 반복 실험과 EKF + LSTM 잔차 보정(Hybrid Residual) 구조에 대한 소거 비교도 함께 들어간다.

본 연구에서는 OOD 조건에서 LSTM 기반 자세추정이 EKF를 보완할 수 있는지 확인하기 위하여 두 가지 가설을 설정하였다. 첫 번째 가설은, 바이어스와 드리프트가 더해진 L2와 L3에서 EKF의 RMSE가 L0·L1 대비 뚜렷하게 증가하여 두 구간이 학습 분포를 벗어난 고노이즈 평가 영역으로 기능할 것이라는 점이다. 이는 EKF의 일반적인 한계를 알아보려는 것이 아니라, L2와 L3가 OOD 평가 구간으로 충분한지를 확인하려는 것이다.

두 번째 가설은 LSTM의 OOD 보완 성능에 관한 것이다. LSTM은 저노이즈 구간인 L0과 L1 데이터로만 학습하고, 학습에 사용하지 않은 L2와 L3에서 EKF와 비교한다. 이때 LSTM이 EKF보다 낮은 RMSE를 보인다면, IMU와 지자기센서의 시계열 패턴을 학습한 모델이 고노이즈 조건에서 EKF의 자세추정 오차를 일부 보완할 수 있음을 의미한다.

본 연구는 L0과 L1로 학습한 LSTM이 학습에 포함되지 않은 L2와 L3 조건에서 EKF의 평균 자세추정 오차를 보완할 수 있는지를 검증한다.

논문의 구성은 다음과 같다. 2장에서는 선행연구 검토와 이론적 배경을 살펴본다. 3장에서는 제안하는 연구 방법을 기술하고, 4장에서는 실험 환경과 학습 설정에 대해 정리한다. 5장에서는 실험 결과와 분석을 다루며, 6장에서 결론과 한계점, 향후 연구 방향을 정리한다.

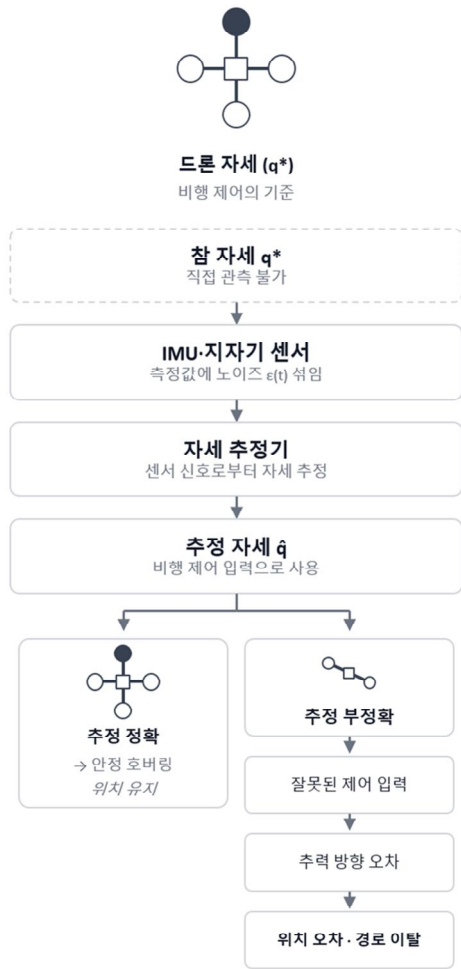


Fig. 1. Effect of drone attitude estimation error on flight control and position stability

II. Related Works

1. Quaternion-Based Attitude Representation

자세를 수학적으로 표현하는 방식에는 오일러 각, 회전 행렬, 쿼터니언이 있다. 오일러 각은 직관적이지만 피치 각이 $\pm 90^\circ$ 가 될 때 짐벌 락이 발생하는 문제가 있고, 회전 행렬은 9개 원소의 직교 조건을 유지해야 하므로 연산에 부담이 있다. 쿼터니언은 네 개의 실수로 자세를 기술하는 방식으로, 짐벌 락이 없고 회전 합성을 간결하게 처리할 수 있으며 단위 노름 조건만 유지하면 되기 때문에 수치 정규화도 단순하다. 이런 특성 덕분에 쿼터니언은 항공·항법 분야 자세추정에서 널리 쓰여 왔다[3, 4]. ArduPilot과 PX4 같은 오픈소스 비행 컨트롤러도 내부 자세 표현으로 쿼터니언을 사용하고, 본 연구에서도 EKF의 자세 추정값과 LSTM의 예측값을 모두 쿼터니언 형식으로 비교한다.

다만 쿼터니언에는 부호 모호성이라는 특성이 있다. q 와

$-q$ 가 같은 회전을 나타낸다는 점이다. 학습 기반 모델이 이를 반영하지 않은 채 평균 제곱 오차를 그대로 쓰면, 같은 자세를 가리키는 두 부호가 서로 다른 값으로 해석되어 학습이 불안정해진다. 본 연구에서는 내적의 절댓값에 기반한 손실 함수로 이 문제를 처리하였으며, 구체적인 정의는 3장 2절에서 다룬다. 두 쿼터니언 사이의 각도 오차도 마찬가지로 내적 절댓값에 역 코사인을 적용해 산출한다. 두 자세가 같으면 오차는 0이고, 두 회전 사이의 상대 회전각이 180° 에 가까워질수록 오차는 최댓값에 가까워진다.

2. EKF-Based Sensor Fusion and Noise-Model Assumptions

노이즈가 섞인 측정값으로부터 시스템 내부 상태를 추정하는 기법으로 1960년 Kalman이 처음 제안한 것이 칼만 필터(Kalman Filter)이다[10]. 예측과 갱신을 반복하며 상태 추정치를 갱신해 나가는 방식이며, 표준형은 선형 시스템과 가우시안 노이즈를 전제로 설계된다.

비선형 문제인 자세추정에는 표준 칼만 필터를 그대로 쓰기 어렵다. 이를 위한 변형이 확장 칼만 필터(Extended Kalman Filter, EKF)이다. 비선형 상태 전이 함수와 측정 함수를 현재 추정치 부근에서 1차 테일러 전개로 근사한 뒤 예측과 갱신을 수행한다[9]. 쿼터니언 기반 EKF에서는 상태 벡터 안에 쿼터니언이 포함되며, 갱신 단계 이후에는 단위 노름 조건을 유지하기 위한 정규화 과정이 따라 붙는다[3, 4]. UAV 상태추정 연구에서 EKF 기반 센서 융합은 폭넓게 사용되어 왔고[4], 본 연구에서도 ArduPilot SITL이 제공하는 자세 추정 출력을 EKF 기준선으로 활용한다.

EKF 기반 자세추정에는 몇 가지 제약이 따른다. 우선 추정 정확도가 프로세스 노이즈 공분산(Q)과 측정 노이즈 공분산(R) 튜닝에 민감하게 좌우된다[7, 11]. Q 와 R 값은 보통 특정 비행 조건에서 반복 실험을 거쳐 맞추는데, 센서 특성이나 비행 환경이 바뀌면 같은 값을 그대로 쓰기 어렵다. 가속도계나 지자기센서가 출력하는 값에는 자세 정보뿐 아니라 외부 가속이나 자기교란 성분까지 함께 섞여 들어가므로, 측정값만으로 자세 성분과 교란 성분을 분리하는 일도 쉽지 않다[5, 6].

비선형성이 강해지거나 노이즈 분포가 설계 시 가정된 범위를 벗어나면 선형화 오차와 모델 가정의 괴리가 동시에 커진다[9]. 자기교란 환경에서 기존 필터의 방위각 오차가 10° 이상까지 벌어진 사례가 보고되어 있다[6]. 가정이 흔들리는 조건에서는 EKF만으로 강건한 추정을 유지하기가 쉽지 않다.

3. LSTM and Time-Series Attitude Estimation

은닉층의 순환 연결을 통해 과거 시점 정보를 현재 시점으로 넘겨주는 구조의 신경망이 순환신경망(Recurrent Neural Network, RNN)이며, 시계열 처리에 자주 활용된다[7]. 다만 기본 RNN에는 한 가지 한계가 있다. 역전파 과정에서 시간 단계가 길어질수록 기울기가 0으로 수렴하거나 발산하는 경향이 나타나, 수백 스텝 떨어진 과거 정보를 현재 출력에 반영하기 어렵다[8].

기울기 소실 문제를 풀고자 등장한 RNN의 변형이 LSTM이다. 셀 상태라는 별도의 정보 저장 경로, 그리고 이를 제어하는 망각·입력·출력 게이트가 LSTM의 핵심을 이룬다[8]. 망각 게이트는 셀 상태에 들어 있는 과거 정보 중 무엇을 유지할지를 결정한다. 새로운 정보를 얼마만큼 반영할지는 입력 게이트가 조절한다. 출력 게이트는 갱신된 셀 상태에서부터 은닉 상태를 산출한다. 셀 상태가 별도 경로로 정보를 전달하는 덕분에 기울기가 여러 시간 단계를 거쳐도 잘 소실되지 않으며, 그 결과 장기 의존성 학습이 가능해진다.

최근 들어 IMU 기반 자세추정에 LSTM을 적용한 연구가 늘고 있다. 6축 IMU 입력으로 기울기 자세를 추정하는 LSTM 모델은 가속 조건에서 상보필터(평균 RMSE 5.16°)와 칼만필터(2.46°)에 비해 평균 1.94°의 오차를 달성한 바 있다[5]. 9축 IMU를 입력으로 받아 쿼터니언 형태의 3차원 자세를 종단간 추정하는 순환신경망의 경우, 자기교란 환경에서 상보필터 대비 41.4%, 칼만필터 대비 60.3% 낮은 RMSE가 보고되었다[6]. 운용 중에도 LSTM 가중치를 점진적으로 갱신하는 증분학습 방식으로 장기 운용 안정성을 끌어올린 연구도 있다[7]. 딥러닝 기반 IMU 자세추정의 가능성은 국내 연구에서도 정리된 바 있으며, 학습 기반 접근의 방향성이 함께 논의되었다[11].

이상 살펴본 선행 연구는 학습 분포 안에서의 성능 향상을 주로 다루었다는 한계가 있다. 학습 분포를 벗어난 OOD 노이즈 조건에서도 데이터 기반 접근이 강건한지에 대한 검증은 충분히 다루어지지 않았기에 본 연구에서 살펴보고자 한다.

4. Out-of-Distribution Evaluation

기계학습 모델 평가에서 가장 흔히 쓰이는 방식은 전체 데이터를 무작위로 학습·검증·테스트 세트로 나누는 것이지만 자세추정 문제에서는 노이즈 수준이 서로 달라 이 방식을 적용하기는 어렵다. L0-L3을 무작위로 분할하면 고노이즈 구간 데이터 일부가 학습 세트에 포함되어 일반화 능력이 실제보다 높게 측정되는 결과가 나올 수 있기 때문

이다. OOD 분할은 학습에 쓰이지 않은 조건의 데이터만을 평가에 사용하여 고노이즈 조건에 대한 성능이 과대평가되는 문제를 방지한다. 학습 분포를 벗어난 입력에 모델이 어떻게 반응하는지를 직접 확인할 수 있다. 학습기반 모델의 일반화 검증은 다중 센서 융합 연구에서도 후속 과제로 짚어진 바 있다[9]. 본 연구에서는 노이즈 레벨을 기준으로 OOD 분할을 적용하여 L0-L1만으로 모델을 학습하고, L2-L3은 평가 전용으로 구분한다.

III. Proposed Method

본 연구가 제안하는 방법은 노이즈 레벨 기반의 OOD 분할과 쿼터니언 출력 LSTM 학습, 그리고 분리된 평가 구간에서의 비교 실험으로 구성된다. 전체 흐름은 Fig. 2와 같이 데이터 준비 단계와 학습·평가 단계로 나뉜다.

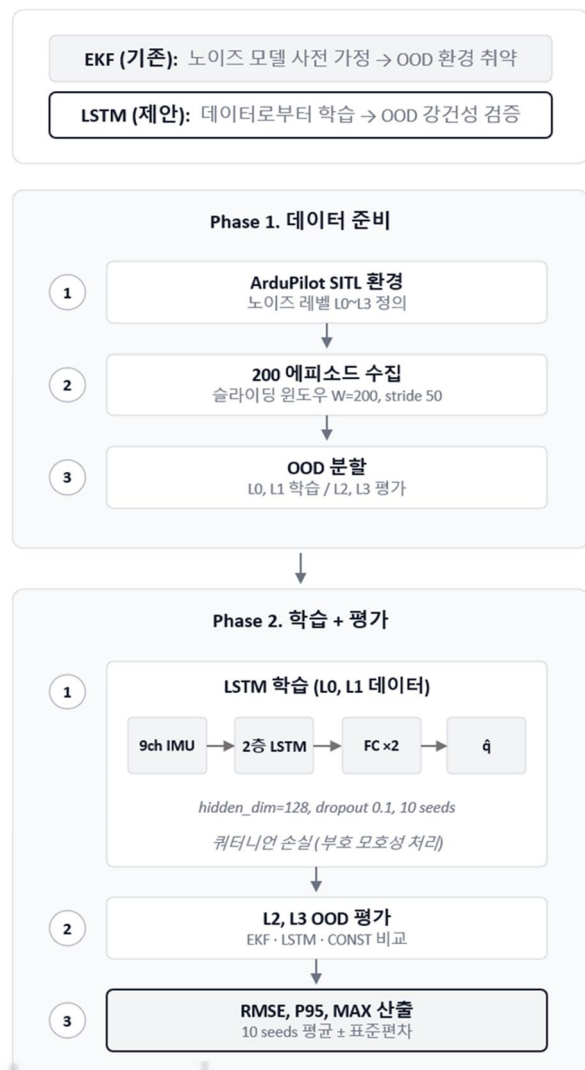


Fig. 2. Overall flow of the proposed method

1. Noise Level Design

OOD 평가를 위해서는 학습 구간과 평가 구간을 분리할 기준이 필요하므로 노이즈 레벨이 먼저 정의되어야 한다. 이에 따라 ArduPilot SITL이 제공하는 시뮬레이션 파라미터를 활용하여 단계별로 레벨을 설계하였다. 가속도계, 자이로스코프, 자력계의 센서 노이즈(SIM_ACC1_RND, SIM_GYR1_RND, SIM_MAG_RND), 자이로 바이어스 3축(SIM_GYR1_BIAS_X/Y/Z), 드리프트속도(SIM_DRIFT_SPEED)의 파라미터를 사용하였다. 이를 L0(노이즈 없음)부터 L3(극한 노이즈)까지 네 단계로 나누었고, 레벨별로 정의한 값은 Table 1에 정리하였다.

Table 1. Noise parameter settings for each level

Level	SIM_ACC1_RND	SIM_GYR1_RND	SIM_MAG_RND	SIM_GYR1_BIAS_X/Y/Z	SIM_DRIFT_SPEED
L0	0.0	0.0	0.0	0.00/ 0.00/ 0.00	0.00
L1	1.0	0.5	2.0	0.00/ 0.00/ 0.00	0.00
L2	3.0	2.0	10.0	0.02/ 0.02/ 0.02	0.20
L3	10.0	5.0	10.0	0.05/ 0.05/ 0.05	0.50

*All levels use ARMING_CHECK = 0 and SIM_DRIFT_TIME = 5.0.

L0과 L1은 학습용으로 바이어스와 드리프트를 0으로 고정하였다. L0은 노이즈를 추가하지 않았고, L1에는 낮은 수치의 랜덤 노이즈를 더했다. L2와 L3은 평가용으로 자이로 바이어스에 측당 0.02 rad/s 이상의 값을 추가하고, 드리프트가 시간이 지남에 따라 누적되도록 하였다. 평가용 레벨에 파라미터값을 추가한 이유는 바이어스나 드리프트와 같은 노이즈에 모델이 어떻게 반응하는지를 확인하려는 것이다.

평가 구간 안에서도 난이도를 달리하였다. L0-L3은 특정 IMU 제품의 데이터시트 값을 그대로 재현한 것이 아니라, UAV 운용 중 발생하는 센서 열화를 단계별로 모사하기 위한 설정이다. L2는 저가형 MEMS IMU 수준의 조건을 가정하였다. 중간 정도의 바이어스와 드리프트에 약한 자기교란을 더한 구간이다. L3은 여기에 가속도계와 자이로스코프의 랜덤 노이즈를 크게 키우고 바이어스와 드리프트를 강화하였다. 모터 진동이 심하거나 온도가 변하고 주변 전자장품의 자기교란이 겹치는, 여러 교란이 한꺼번에 커지는 상황을 상정한 것이다. 노이즈 강도가 높아질수록

모델 오차가 어떻게 변하는지 보기 위해 두 단계로 구분하였다.

2. LSTM Model for Attitude Estimation

OOD 평가의 대상이 되는 LSTM 모델은 본 절에서 기술한다. 시계열 IMU 신호로부터 쿼터니언 자세를 직접 출력하는 종단간(end-to-end) 구조이며, 쿼터니언 부호 모호성을 처리할 수 있는 손실 함수와 함께 동작한다.

2.1. Input Window and Channel Configuration

LSTM 모델에 들어가는 입력은 SI 단위로 변환된 9축 IMU 신호이다. 가속도 3축, 각속도 3축, 자기장 3축으로 구성된 9개 채널이 시계열 입력 하나를 이룬다. 시계열 입력은 슬라이딩 윈도우 방식으로 만들어진다. 윈도우 크기 $W = 200$ 샘플(100 Hz 기준 2초)에 스트라이드 50 샘플을 적용했으므로, 윈도우는 0.5초 간격으로 이동한다. 한 에피소드의 평균 행수가 약 2,904이라는 점을 감안하면 에피소드당 약 55개의 윈도우가 생성된다. 각 윈도우의 형태는 (200, 9)이며, 모델은 이 윈도우를 받아 마지막 시점에 대응하는 쿼터니언 하나를 출력한다.

2.2. Model Architecture

9채널 IMU 윈도우 (200, 9)를 입력으로 받은 모델은 hidden_dim 128의 2층 단방향 LSTM(dropout 0.1)을 거치고, 두 단의 fully-connected 층(128 → 64, ReLU 활성화 / 64 → 4)을 통과한다. 마지막 단계에서는 L2 정규화로 단위 노름 조건($\|q\| = 1$)을 보장해 4차원 쿼터니언을 출력한다.

은닉 차원을 128로 설정한 이유는 9채널 입력의 시간 패턴을 학습하기에 충분하면서도 과적합을 피할 수 있는 수준이기 때문이다. 표현력을 확보하되, 깊이가 과도해질 때 발생하는 학습 불안정을 방지하고자 하였다. 양방향 대신 단방향 구조를 택한 이유는 실시간 추론 환경에서 미래 시점 데이터를 참조할 수 없다는 데 있다. 출력단에는 L2 정규화를 두어 단위 쿼터니언 조건을 자동으로 만족하도록 하였으며, 별도의 후처리 없이도 유효한 자세 표현이 출력된다.

2.3. Loss Function

손실 함수는 두 쿼터니언 사이의 각도 차이를 반영하도록 다음과 같이 정의한다.

$$\text{QuaternionLoss} = 1 - |q_{\text{pred}} \cdot q_{\text{gt}}|$$

q_{pred} 는 모델이 출력한 단위 쿼터니언이고, q_{gt} 는 SITL 로그에서 추출한 정답 쿼터니언이다. 두 쿼터니언

사이 각도가 0°일 때 손실은 0이고, 각도가 커질수록 1에 가까워진다. 내적에 절댓값을 씌운 형태는 쿼터니언의 부호 모호성($q \equiv -q$)에 대응하기 위한 것으로, 부호가 반전되어도 손실은 0으로 유지된다.

3. Evaluation Protocol

3.1. Data Split

데이터는 노이즈 레벨을 기준으로 분할 하였다. L0과 L1의 100개 에피소드를 학습에 사용하여 80%는 훈련 20%는 검증에 사용하였고, 검증 세트는 모델 선택과 조기 종료 판단에만 활용된다. L2와 L3의 100개 에피소드는 OOD 테스트에만 사용한다. 입력 정규화에는 Z-score를 적용하되, 학습 세트에서만 산출한 평균과 표준편차를 학습/검증/테스트에 같이 적용한다. 테스트 데이터의 통계가 정규화에 섞여 들어가는 정보 누출을 막기 위함이다.

3.2. Multi-Seed Repetition

42부터 51까지의 시드를 동일한 하이퍼파라미터로 학습하였다. 시드가 달라지면 모델 초기 가중치, 미니배치 구성 순서, 레벨 내 에피소드의 훈련/검증 배분이 함께 달라진다. 시드별 결과의 차이가 모델 성능 평가에 영향을 주지 않도록 10회 결과의 평균과 표준편차를 함께 확인하였다. 시드의 평균은 성능의 대푯값이고 표준편차는 시드 간 변동 폭을 나타낸다.

3.3. Evaluation Metrics

자세추정 성능은 2장 1절에서 기술한 각도 오차 정의를 바탕으로 세 가지 지표로 평가한다. RMSE는 추정 정확도의 전반적 수준을 나타내고, P95는 상위 구간의 오차 경향을 함께 반영하며, MAX는 모델이 순간적으로 보이는 최악의 편차를 가리킨다. RMSE에 의존하면 짧은 시간 동안 발생하는 큰 편차가 평균에 묻힐 수 있고, MAX만으로는 전반적인 추정 품질을 가늠하기 어렵다. P95는 두 지표를 보완하며 상위 오차 경향을 보여준다. 세 지표는 각자 다른 측면을 반영하므로 모델의 강건성을 다방면으로 평가하기 위해 세 지표를 함께 확인하였다.

3.4. Baseline Models

비교 대상으로 두는 모델은 세 가지이다. 그중 첫째인 ArduPilot EKF의 출력은 ATTITUDE_QUATERNION 메시지로 수신되며, IMU와 지자기센서를 포함한 다중 센서 융합을 거쳐 산출된다[4]. 본 연구의 주된 비교 대상이기도 하다. 둘째 비교 대상은 앞서 기술한 LSTM 모델이며, OOD 조건에서 EKF 대비 오차 감소가 핵심 평가 목표가

된다. 마지막은 CONST(1, 0, 0, 0) 기준선으로, 모든 시점에서 항등 쿼터니언을 출력한다. 아무 추정 행위 없이 나올 수 있는 오차의 상한을 보여주는 역할이며, EKF와 LSTM이 의미 있는 추정을 수행하는지 점검한다.

4. Ablation Study Design

소거 비교에서는 윈도우 크기 효과를 보기 위한 W=100과 W=200의 LSTM, 순환 구조 비교를 위한 GRU(W=200), 그리고 EKF와 LSTM을 결합한 Hybrid Residual(W=200)을 살펴본다. 윈도우 비교와 GRU 비교는 LSTM 단독 모델의 설계 선택지(시계열 길이, 게이트 구조)가 OOD 성능에 어떤 영향을 주는지를 점검하기 위한 것이고, Hybrid Residual은 EKF의 안정성과 LSTM의 OOD 적응력을 함께 활용할 수 있는지 확인하기 위한 구조이다.

신경망을 센서 오차 보정이나 기존 추정 파이프라인의 보완 모듈로 쓰는 접근은 선행연구에서도 보고된 바 있다 [12]. 본 연구는 이 아이디어를 EKF 출력 자세에 대한 쿼터니언 잔차 보정 구조로 확장한다.

IV. Experimental Setup

1. Simulation Environment

실험 데이터는 Ubuntu 22.04 환경에서 ArduPilot SITL을 구동해 수집하였다. 실제 하드웨어 없이 비행 컨트롤러 소프트웨어를 가상으로 실행하는 도구가 SITL이며, 내부 센서 모델과 물리 엔진을 통해 다양한 센서 조건을 재현해 준다. SITL을 사용한 이유는 노이즈 파라미터를 직접 제어할 수 있어 OOD 평가에 필요한 노이즈 분포 차이를 명확히 두기 쉽고, 실기체 실험에서의 환경 변동을 배제한 채 노이즈 외 변수를 고정한 상태에서 노이즈 영향만 관찰할 수 있기 때문이다.

비행 시나리오는 호버링으로 한정하였다. 비행 동작에 따른 자세 변화가 작아 노이즈의 영향을 분리해 관찰하기에 알맞기 때문이다. 웨이포인트 추종이나 원형 비행 같은 동적 시나리오에서는 비행 동작 자체의 자세 변화와 노이즈 효과가 함께 작용하므로, 노이즈의 영향만을 보기 어렵다.

시뮬레이터와의 통신에는 MAVLink 프로토콜과 pymavlink 라이브러리를 사용하였다. 센서 데이터, Ground Truth, EKF 추정값은 세 종류의 MAVLink 메시지로 수집한다. 100Hz 주기로 9축 IMU 측정값을 전달하는 것이 RAW_IMU이고, 20Hz 주기로 시뮬레이터 내부 물리 엔진이 산출한 참값 쿼터니언을 전달하는 것이

SIM_STATE이다. ATTITUDE_QUATERNION은 20Hz 주기로 ArduPilot 내장 EKF의 추정 자세를 전달한다.

메시지들은 샘플링 레이트가 다르기 때문에 이 데이터를 동기화하기 위해 RAW_IMU의 메시지가 도착할 때, SIM_STATE와 ATTITUDE_QUATERNION의 가장 최근 값을 그대로 가져오는 hold 방식을 사용하였다. 동기화 지연을 차후에 확인할 수 있도록, 각 행에는 IMU·GT·EKF 타임스탬프를 따로 기록해 두었다.

2. Episode-Based Data Collection

데이터 수집은 L0부터 L3까지 레벨당 50번씩 총 200개의 에피소드를 확보하였다. 각 에피소드는 SITL 프로세스를 새로 시작하고 해당 레벨의 노이즈 파라미터를 설정한 뒤에 실행된다. 드론을 이륙시키고 호버링 자세가 안정될 때까지 약 20초간 대기 후 EKF 내부 상태가 수렴하도록 15초의 burn-in 구간을 두었다. 그 후 30초 동안 100Hz로 센서 데이터를 수집하였다. 통신 지연 등의 영향으로 실제 수집된 행수는 에피소드당 평균 약 2,904행으로 나타났다.

에피소드마다 SITL을 완전히 재시작한 이유는 EKF 내부 상태를 초기화해 회차 간 독립성을 유지하려는 데 있다. ArduPilot의 EKF는 내부 상태를 유지한 채 갱신되는 구조이기 때문에, 에피소드를 연속으로 실행하면 직전 회차가 다음 회차의 초기 조건에 영향을 줄 수 있다.

분석에 포함되는 에피소드는 네 가지 기준을 통과해야 한다. 결측 비율 5% 이하, 행 수 2,550 이상(30초 × 100Hz 기대치의 85%), 타임스탬프 단조성, watchdog 미발동의 네 가지이다. 200개 모두 기준을 통과해 별도 제외 없이 분석에 사용하였다.

3. Training Configuration

Table 2는 학습에 사용한 하이퍼파라미터를 정리한 것이다.

Table 2. Training hyperparameters

Item	Value
Training levels	L0, L1
Test levels	L2, L3
Validation ratio	0.2
Maximum epochs	50
Batch size	64
Optimizer	Adam
Initial learning rate	0.001
Weight decay	1e-5
LR scheduler	ReduceLROnPlateau (patience=5, factor = 0.5)
Early stopping	patience = 15
Random seeds	42-51 (10 seeds)

옵티마이저는 시계열 학습에서 자주 쓰이는 표준 방식인 Adam을 사용하였다. 학습률 스케줄러는 ReduceLROnPlateau를 사용하였는데, 검증 손실이 5 에폭 연속 나아지지 않을 때 학습률을 절반으로 줄이는 방식이다. 학습률은 0.001에서 0.0005, 0.00025, 0.000125 순으로 낮아진다. 조기 종료의 작동 조건은 검증 손실이 15 에폭 동안 개선되지 않을 때이며, 검증 손실이 가장 낮았던 시점의 모델 가중치가 최종 평가에 활용된다. 최대 에폭 수는 50으로 두었고, 4개의 시드에서는 조기 종료가 작동하였다.

V. Experimental Results

1. Noise Injection Verification

3장 1절에서 정의한 노이즈 레벨 설계가 의도대로 작동하는지를 실험에 앞서 두 가지 방식으로 확인하였다.

먼저 가속도계(SIM_ACC1_RND), 자이로스코프(SIM_GYR1_RND), 자력계(SIM_MAG_RND)의 노이즈 파라미터를 각각 단계적으로 변화시키며 센서 출력 분산을 측정하였다. 가속도계와 자이로스코프는 레벨이 높아질수록 분산이 뚜렷이 커지면서 노이즈 파라미터가 출력에 의도대로 반영됨을 확인하였다. 자력계는 L2와 L3의 SIM_MAG_RND 값이 같지만 L3에서 분산이 더 크게 나타났다는데, L3의 강한 자이로 드리프트가 기체 자세 변동을 키워 자력계가 측정하는 자기장 벡터의 변동 폭도 함께 늘어난 결과로 해석된다.

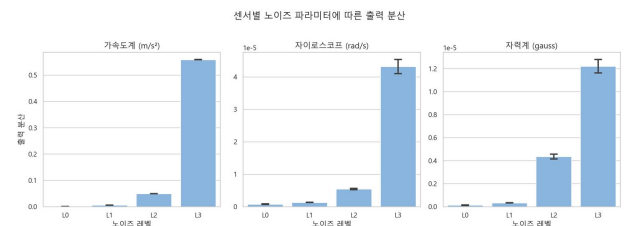


Fig. 3. Sensor output variance with respect to noise parameters

복합 노이즈 레벨이 EKF 자세추정 오차에 미치는 영향은 200개 에피소드에서 측정하였다.

Table 3. EKF RMSE statistics by level (deg)

Level	Episodes	RMSE Mean	RMSE Std
L0	50	0.8252	0.0296
L1	50	0.8143	0.0213
L2	50	2.3918	0.0321
L3	50	4.8079	0.1546

Table 3은 노이즈 레벨별 EKF RMSE를 보여준다. L0과 L1의 RMSE는 각각 0.8252° , 0.8143° 로 1° 미만이고, 0.011° 의 차이는 표본 변동 범위 안에 들어간다. RMSE는 L2에서 2.3918° 로, L3에서는 4.8079° 까지 늘어났다. L2 이상에서는 바이어스와 드리프트가 추가되기 때문에 오차가 함께 늘어났고, L3의 표준편차가 0.1546° 까지 커진 점도 같은 흐름에서 읽을 수 있다.

L2 이상에서 EKF의 RMSE가 증가하였고, L3에서는 에피소드 간 변동 폭도 커졌다. 이는 L2와 L3를 고노이즈 OOD 평가 구간으로 분리한 설정을 뒷받침하며, 서론에서 제기한 첫 번째 가설과도 일치한다.

Table 3의 EKF RMSE는 에피소드 전체 시계열에서 계산한 평균값이다. 다음 절 Table 4의 EKF RMSE는 LSTM 평가 시점에 맞춘 슬라이딩 윈도우 단위로 계산된 값이다. 두 표 사이의 EKF 수치 차이는 집계 단위가 다른 데에서 비롯된다.

2. EKF vs. LSTM Performance

본 절에서는 제안하는 LSTM 모델의 전체 성능을 EKF 및 CONST 기준선과 비교한다. Table 4는 학습에 사용되지 않은 L2·L3에서 EKF와 LSTM의 OOD 성능을 정리한 것이다.

Table 4. OOD test performance (deg, mean $\pm$ std over 10 seeds)

Level	Model	RMSE	P95	MAX
L2	CONST (1,0,0,0)	7.7603 ± 0.0000	8.2559 ± 0.0000	10.6036 ± 0.0000
L2	EKF	2.4249 ± 0.0000	2.9100 ± 0.0000	2.9924 ± 0.0000
L2	LSTM	1.9062 ± 1.1289	2.3283 ± 1.1335	4.3924 ± 0.6800
L3	CONST (1,0,0,0)	10.5912 ± 0.0000	12.0711 ± 0.0000	16.4541 ± 0.0000
L3	EKF	4.9075 ± 0.0000	6.3389 ± 0.0000	6.5952 ± 0.0000
L3	LSTM	4.2133 ± 0.5316	5.4352 ± 0.4585	9.7485 ± 0.3603

EKF는 같은 입력에 대해 항상 같은 출력을 산출하고, CONST는 고정된 항등 쿼터니언을 출력하므로 두 모델 모두 시드 간 표준편차가 0이다. LSTM의 표준편차는 10개 랜덤 시드 간 변동을 의미한다. MAX는 OOD 테스트 세트에서 발생한 각도 오차 중 가장 큰 값을 의미하며, 각 시드 별로 산출하였다.

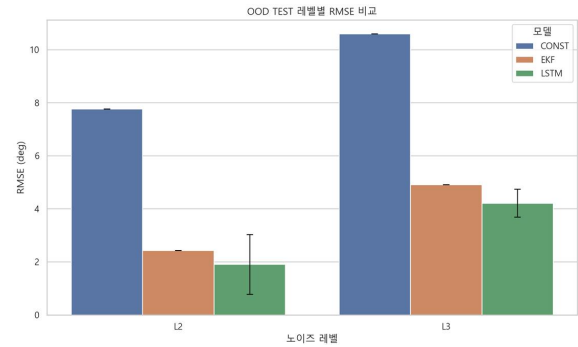


Fig. 4. RMSE comparison on OOD test (EKF vs. LSTM vs. CONST, L2/L3)

학습 분포 내(L0, L1) 검증 성능을 살펴보면, L1에서 LSTM RMSE는 $0.1082^\circ \pm 0.0218^\circ$ 로 EKF($0.7982^\circ \pm 0.0043^\circ$) 대비 86.4% 낮았다. L0에서도 LSTM 평균 (0.2565°)이 EKF(0.8106°)보다 낮았다. 다만 일부 시드의 L0 검증 RMSE가 다른 시드보다 높게 나오면서 L0 LSTM 표준편차가 0.2460° 까지 벌어졌다. 이 현상은 5.3절 시드 민감성 분석에서 다시 다룬다.

OOD 구간에서도 평균 오차에서 LSTM이 EKF보다 낮은 RMSE를 기록하였다. RMSE를 기준으로 LSTM이 EKF보다 L2에서 21.4%, L3에서 14.1% 낮았다. P95에서도 같은 흐름이다. L2에서 20.0%, L3에서 14.3% 감소하였다. 에피소드 단위로 확인해 보면 두 레벨 모두 약 88%의 에피소드에서 LSTM의 RMSE가 EKF보다 낮게 측정되었다. CONST 모델은 두 모델보다 훨씬 큰 값을 기록하였는데, L2에서 7.7603° , L3에서 10.5912° 였다. EKF와 LSTM 모두 자세추정에 유효하게 작동하였다. 이 결과는 서론에서 설정한 두 번째 가설과 부합한다. L0과 L1만으로 학습한 LSTM이 학습에 포함되지 않은 L2와 L3 조건에서도 RMSE 기준으로 EKF보다 낮은 오차를 보였기 때문이다.

다만 MAX 지표에서는 반대 양상이 나타났다. L2에서는 EKF보다 LSTM이 약 1.40° 높았다. L3에서는 격차가 더 벌어져 EKF 6.5952° , LSTM 9.7485° 로 3.15° 차이를 보였다. 대부분의 시점에서는 LSTM이 더 정확하지만, 일부 윈도우에서 순간적으로 큰 오차가 발생한다는 뜻이다. 물리 모델 기반으로 동작하는 EKF는 출력 범위가 안정적인 반면, 데이터에서 학습한 LSTM은 OOD 입력의 일부 시점에서 큰 오차를 보일 수 있다. 안전 관련 응용에서 MAX 오차는 가볍게 보기 어려운 지표이므로, 그 완화 가능성은 5.3절 소거 실험의 Hybrid Residual 결과와 함께 다룬다.

3. Ablation Study

Table 5는 3장 4절에서 설계한 Hybrid Residual 모델을 포함하여 네 가지 변형 모델의 소거 실험 결과를 비교한 것이다. 이를 통해 윈도우 길이, 순환 구조, Hybrid Residual 결합 방식이 오차에 미치는 영향을 살펴보았다. 각 변형 모델은 계산 비용을 고려해 시드 3개(42, 43, 44)로만 반복 수행하였다. 본 절에서는 Ablation 결과와 함께 앞 절 주 실험(시드 10개)에서 관찰된 시드 민감성도 정리한다.

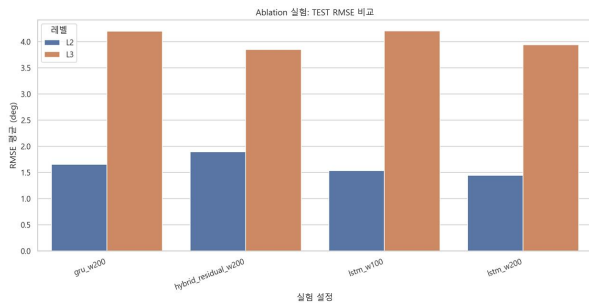


Fig. 5. Ablation RMSE comparison

Table 5. Ablation results (deg, mean $\pm$ std over 3 seeds)

Variant	Level	RMSE	P95	MAX
LSTM W=100	L2	1.5361 $\pm$ 0.0639	1.9788 $\pm$ 0.0414	4.2191 $\pm$ 0.0505
LSTM W=200	L2	1.4461 $\pm$ 0.2047	1.8954 $\pm$ 0.1872	4.0621 $\pm$ 0.3338
GRU W=200	L2	1.6589 $\pm$ 0.1726	2.0478 $\pm$ 0.1484	15.0027 $\pm$ 9.7723
Hybrid Residual W=200	L2	1.8951 $\pm$ 0.6104	2.0580 $\pm$ 0.2278	20.9773 $\pm$ 26.4307
LSTM W=100	L3	4.2087 $\pm$ 0.1532	5.4931 $\pm$ 0.1751	9.9152 $\pm$ 0.3365
LSTM W=200	L3	3.9438 $\pm$ 0.4557	5.1741 $\pm$ 0.5316	9.5565 $\pm$ 0.5561
GRU W=200	L3	4.2002 $\pm$ 0.3808	5.5166 $\pm$ 0.4013	10.2321 $\pm$ 0.3460
Hybrid Residual W=200	L3	3.8501 $\pm$ 0.1293	5.2237 $\pm$ 0.1548	5.8165 $\pm$ 0.3748

3.1. Window Size Effect (W=100 vs. W=200)

LSTM W=200의 RMSE는 L2에서 1.4461° , L3에서 3.9438° 로 나타났다. W=100(1.5361° , 4.2087°)과 비교하면 5.9~6.3% 낮은 수치이다. 윈도우 길이가 1초에서 2초로 늘어난다면 모델이 참조하는 시계열 구간도 함께 길어지므로, 노이즈 영향을 줄이는 데 유리해진다.

3.2. LSTM vs. GRU

같은 W=200 조건에서 LSTM의 RMSE는 GRU보다 L2에서 12.8%, L3에서 6.1% 낮았다. 게이트 구조가 단순해

파라미터 수가 적다는 점은 GRU의 장점이다. 다만 자세추정처럼 장기 시간 의존성이 중요한 문제에서는 LSTM의 셀 상태가 더 효과적으로 작동한 것으로 해석된다. GRU의 L2 MAX는 $15.0027 \pm 9.7723^\circ$ 로 표준편차가 컸으며, 일부 시드에서 큰 MAX 값이 발생하여 GRU의 시드 간 변동성이 LSTM보다 높음을 보였다.

3.3. Hybrid Residual Model

L3에서 Hybrid Residual의 RMSE는 3.8501° 로 LSTM의 3.9438° 와 0.0937° 격차에 머문다. MAX는 5.8165° 로 LSTM의 9.5565° 보다 39% 낮은 값이다. EKF 출력에 LSTM 잔차를 더하는 결합 방식이 LSTM 단독 모델의 MAX 변동을 누르는 효과를 갖는다. L2에서는 RMSE가 1.8951° 로 LSTM의 1.4461° 보다 0.4490° 높게 측정되었다. MAX는 $20.9773 \pm 26.4307^\circ$ 로 표준편차가 평균을 넘었다. 이는 L3에서 Hybrid Residual이 MAX를 낮춘 것과 다른 양상이다. L3는 바이어스와 드리프트가 강해 EKF 오차가 누적되는 경향이 상대적으로 크므로, LSTM 잔차가 이를 보정하는 방향으로 작동한 것으로 보인다. 반면 L2는 EKF 기본 오차가 작아 잔차 보정 대상이 약한 구간이며, 이 경우 잔차 보정이 일부 윈도우에서 과보정으로 이어졌을 가능성이 있다. 표준편차가 평균을 넘는 점으로 보아, 전체 구간의 일관된 저하라기보다 일부 시드나 윈도우의 순간 오차가 주도한 결과로 해석된다.

3.4. Seed Sensitivity Analysis

주 실험의 10개 시드 중 시드 51이 이상치로 나타났다. 시드 42에서 50까지의 L2 RMSE는 1.27° 에서 1.98° 사이에 분포하였으나, 시드 51만 5.25° 를 기록해 다른 9개 시드 평균(1.53°)의 3배 이상이었다. L3에서도 시드 51(5.63°)이 나머지 평균(4.06°)보다 뚜렷이 높았다.

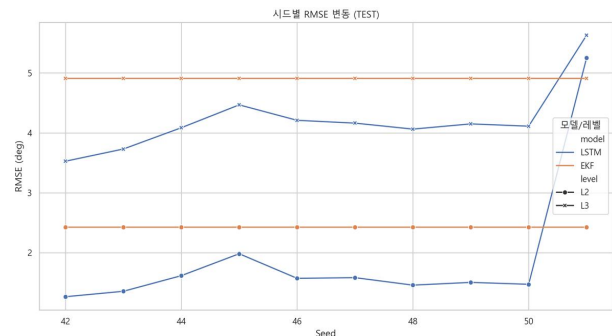


Fig. 6. Per-seed OOD RMSE variation (L2, L3)

시드 51의 L0 검증 RMSE가 0.6525° 로 정상적으로 수렴한 시드들($0.08 \sim 0.12^\circ$)의 5~8배에 달했다는 점을 감안하면, 초기 가중치가 불리한 국소해로 수렴하면서 학습 분포 내 검증 단계부터 불안정했던 것으로 짐작할 수 있다. 시드 51을 제외할 경우 L2 LSTM RMSE 평균은 약 1.53° , 표준편차는 약 0.20° 로 안정된다. 본 연구에서는 시드 51을 포함한 10회 전체 결과를 주요 수치로 보고하되, 시드 51을 이상치로 별도 표시하였다.

VI. Conclusions

본 연구는 저노이즈 구간으로 학습한 LSTM이 학습에 포함되지 않은 고노이즈 OOD 조건에서 EKF를 보완할 수 있는지를 검증하였다. 실험 환경은 ArduPilot SITL을 기반으로 한다. 가속도계, 자이로스코프, 지자기센서의 노이즈에 자이로 바이어스와 드리프트 속도를 결합해 L0부터 L3까지 네 단계의 노이즈 조건을 구성하였고, LSTM은 L0과 L1만으로 학습시켰다. L2와 L3는 OOD 평가 구간으로 따로 분리한다. 비교 대상은 ArduPilot EKF, 쿼터니언 출력 LSTM, CONST 기준선이며, 평가에는 RMSE, P95, MAX 지표와 다중 시드 반복 결과를 활용하였다.

첫 번째 가설은 EKF의 RMSE 변화에서 검증 근거를 얻는다. L0과 L1에서 EKF의 RMSE는 낮은 값에 머물렀고, 바이어스와 드리프트가 부가된 L2 이상에서는 RMSE가 큰 폭으로 증가하였다. 오차 상승은 노이즈 파라미터의 강도 증가만으로 설명되지 않는다. LSTM이 학습한 L0과 L1 구간에는 바이어스와 드리프트가 포함되지 않으며, L2와 L3는 두 오차 요소를 포함하는 OOD 평가 영역에 해당한다.

두 번째 가설은 RMSE 기준에서는 지지되었다. L0과 L1만으로 학습한 LSTM은 학습에 포함되지 않은 L2와 L3에서 EKF보다 낮은 평균 RMSE를 기록하였고, P95 지표에서도 동일한 방향의 결과가 도출되었으나, 일부 원도우에서는 LSTM의 순간 오차가 EKF를 상회하는 수준까지 증가하였다. 이 결과를 LSTM이 모든 지표에서 일관되게 개선된 것으로 해석하기 어려운 이유는, 평균 오차에서는 LSTM이 낮은 값을 보였지만 순간 최대 오차에서는 EKF보다 불안정한 결과가 나타났기 때문이다. 따라서 LSTM은 EKF를 대체할 수 있다는 것이 아니라, 고노이즈 조건에서 EKF 기반 자세추정의 평균 오차를 줄이는 보정 모듈로 보는 것이 더 적절하다.

본 연구의 주요 기여와 의의는 단순한 성능 수치 비교를 넘어, OOD 환경에서 학습 기반 자세추정 모델이 EKF를

보완할 가능성을 검증하였다는 점에 있다. 먼저, ArduPilot SITL 기반의 OOD 평가 절차를 제안하여 학습 기반 자세추정 모델의 일반화 성능을 일관된 기준에서 검증할 수 있는 기반을 마련하였다. 이는 본 연구뿐 아니라 후속 연구에서도 학습 분포를 벗어난 조건에서의 모델 동작을 정량적으로 비교하는 데 활용될 수 있다. 다음으로, 저노이즈 구간으로만 학습한 LSTM이 학습에 포함되지 않은 고노이즈 조건에서도 평균 RMSE를 낮게 유지함을 보임으로써, 데이터 기반 접근이 센서 환경 변화에 대한 강건성을 일정 수준 확보할 수 있음을 확인하였다. 마지막으로, 평균 오차에서는 LSTM이 낮은 RMSE를 보인 반면, 순간 최대 오차에서는 EKF가 더 안정적인 양상을 보였다. 이는 두 접근이 단순한 우열 관계가 아니라, 운용 지표에 따라 서로 다른 역할을 갖는 보완 관계임을 시사한다.

본 연구는 ArduPilot SITL의 호버링 조건을 기준으로 실험을 수행하였다. 이 설정은 노이즈 레벨을 통제하면서 EKF와 LSTM을 비교하기에는 적합하지만, 실제 기체에서 나타나는 모터 진동, 온도 변화, 전자기 간섭, 급격한 기동까지 포함하지는 못한다. 따라서 실기체 비행 로그를 확보하고, 상승, 회전, 경로 추종과 같은 동적 비행 시나리오로 OOD 평가를 확장하는 것이 후속 과제로 남는다. 또한 일부 시드에서 검증 오차가 크게 증가한 결과를 고려하여, 검증 RMSE가 불안정한 모델은 제외하고 여러 시드에서 학습한 모델의 출력을 평균하는 다중 시드 앙상블 방식을 적용할 필요가 있다. Hybrid Residual 구조에서는 LSTM 잔차가 지나치게 크게 반영되지 않도록 보정량의 범위를 제한하는 방법을 검토하고, 이를 EKF의 안정성과 학습 기반 모델의 보정 성능을 함께 활용하는 적응형 Hybrid 자세추정 구조로 확장하고자 한다.

ACKNOWLEDGEMENT

This work was conducted under the guidance of Professor Seok-Won Lee from the Department of Software and Computer Engineering at Ajou University. Seok-Won Lee's work was supported by the Institute of Information and communications Technology Planning and Evaluation (IITP) under the Artificial Intelligence Convergence Innovation Human Resources Development grant funded by Korean Government [Ministry of Science and ICT (MSIT)] under Grant IITP-2026-RS-2023-00255968.

REFERENCES

- [1] M. Półka, S. Ptak, and Ł. Kuziora, "The Use of UAV's for Search and Rescue Operations," *Procedia Engineering*, Vol. 192, pp. 748–752, 2017. DOI: 10.1016/j.proeng.2017.06.129
- [2] S. Omari, P. Gohl, M. Burri, M. Achtelik, and R. Siegwart, "Visual Industrial Inspection using Aerial Robots," *Proceedings of the 3rd International Conference on Applied Robotics for the Power Industry (CARPI)*, pp. 1–5, October 2014. DOI: 10.1109/CARPI.2014.7030056
- [3] A. M. Sabatini, "Quaternion-based Extended Kalman Filter for Determining Orientation by Inertial and Magnetic Sensing," *IEEE Transactions on Biomedical Engineering*, Vol. 53, No. 7, pp. 1346–1356, July 2006. DOI: 10.1109/TBME.2006.875664
- [4] S. P. H. Driessen, N. H. J. Janssen, L. Wang, J. L. Palmer, and H. Nijmeijer, "Experimentally Validated Extended Kalman Filter for UAV State Estimation Using Low-Cost Sensors," *IFAC-PapersOnLine*, Vol. 51, No. 15, pp. 43–48, 2018. DOI: 10.1016/j.ifacol.2018.09.088
- [5] C. Lee, J. Choi, and J. Lee, "Performance Improvement of IMU-based Attitude Estimation on Accelerated Conditions by Applying an LSTM network," *Proceedings of the ICROS Conference*, pp. 465–467, 2022.
- [6] J. Choi and J. Lee, "Recurrent Neural Network for Nine-Axis IMU-Based Orientation Estimation: 3D Orientation Estimation Performance in Disturbed Conditions," *Journal of Institute of Control, Robotics and Systems*, Vol. 28, No. 12, pp. 1216–1223, December 2022. DOI: 10.5302/J.ICROS.2022.22.0127
- [7] P. Narkhede, R. Walambe, S. Poddar, and K. Kotecha, "Incremental Learning of LSTM Framework for Sensor Fusion in Attitude Estimation," *PeerJ Computer Science*, Vol. 7, e662, August 2021. DOI: 10.7717/peerj-cs.662
- [8] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, Vol. 9, No. 8, pp. 1735–1780, November 1997. DOI: 10.1162/neco.1997.9.8.1735
- [9] Y. Zhuang, X. Sun, Y. Li, J. Huai, L. Hua, X. Yang, X. Cao, P. Zhang, Y. Cao, L. Qi, J. Yang, N. El-Bendary, N. El-Sheimy, J. Thompson, and R. Chen, "Multi-sensor Integrated Navigation/Positioning Systems Using Data Fusion: From Analytics-Based to Learning-Based Approaches," *Information Fusion*, Vol. 95, pp. 62–90, July 2023. DOI: 10.1016/j.inffus.2023.01.025
- [10] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Journal of Basic Engineering*, Vol. 82, No. 1, pp. 35–45, March 1960. DOI: 10.1115/1.3662552
- [11] T. Choi, E. Shin, and W. Nam, "Deep Learning-based Pose Estimation using IMU Accelerometer and Gyroscope Data," *Proceedings of the KSME Spring/Autumn Conference*, pp. 166–168, 2023.
- [12] S. Seong and S. Jeong, "A Calibration Method for IMU Attached to a Mobile Robot Using LSTM Network," *Journal of Institute of Control, Robotics and Systems*, Vol. 29, No. 6, pp. 488–494, June 2023. DOI: 10.5302/J.ICROS.2023.23.0034

Authors



deep learning.

Bu-Yeon Kim is currently pursuing the M.S. degree in Information and Communication Engineering at Ajou University, Suwon, South Korea. Her current research interests include UAV attitude estimation, sensor fusion, and



Seok-Won Lee received the B.Sc. degree in computer science from Dongkuk University, Seoul, Republic of Korea, the M.Sc. degree in computer science from the University of Pittsburgh, PA, USA, and the

Ph.D. degree in information technology from George Mason University, Fairfax, VA, USA. Dr. Lee has been a Full Professor with the Department of Software and Computer Engineering, Ajou University, Republic of Korea, since 2012. Previously, he was a Faculty Member with the University of Texas at San Antonio (2011–2012) and the University of North Carolina at Charlotte (2003–2010), and a Visiting Faculty with the University of Nebraska–Lincoln (2010–2011). Prior to his academic career, he was affiliated with SAIC, IBM T. J. Watson Research Center, and George Mason University. He has published more than 250 refereed articles. His research interests include software engineering (ontological requirements engineering and domain modeling), knowledge engineering (knowledge acquisition, machine learning, and knowledge-based systems), and information assurance (software security and privacy). He is a Senior Member of IEEE, ACM, and AAAI.