

Gated recurrent unit (GRU) 신경망을 이용한 적혈구 침강속도 예측

이재진* · 홍현지** · 송재민*** · 염은섭†

Forecasting of erythrocyte sedimentation rate using gated recurrent unit (GRU) neural network

Jaejin Lee*, Hyeonji Hong**, Jae Min Song***, Eunseop Yeom†

Abstract In order to determine erythrocyte sedimentation rate (ESR) indicating acute phase inflammation, a Westergren method has been widely used because it is cheap and easy to be implemented. However, the Westergren method requires quite a long time for 1 hour. In this study, a gated recurrent unit (GRU) neural network was used to reduce measurement time of ESR evaluation. The sedimentation sequences of the erythrocytes were acquired by the camera and data processed through image processing were used as an input data into the neural network models. The performance of a proposed models was evaluated based on mean absolute error. The results show that GRU model provides best accurate prediction than others within 30 minutes.

Key Words : ESR(적혈구 침강 속도), GRU(게이트 순환 유닛), Erythrocyte aggregation(적혈구 응집)

1. 서론

적혈구 침강속도 (Erythrocyte Sedimentation Rate; ESR)는 수직 시험관내 혈액의 적혈구가 1 시간 동안 혈장으로부터 분리되어 침강한 거리를 의미한다. ESR 측정값은 염증으로 인해 증가한 적혈구 응집도(Erythrocyte aggregation)와 연관이 높다.⁽¹⁾

적혈구 막은 음전하를 띠고 있어, 체내에 염증으로 인해 양전하를 띤 급성기반응물질의 증가는 혈액 내에서 적혈구의 응집을 활성화하고, 3차원적 응집체(Rouleaux)를 형성한다. 침강 속도는 입자의 크기에 비례하기 때문에 적혈구가 응집하여 크기가 커짐에 따라 침강 속도는 빨라지므로, 염증에 의한 응집 활성화는 결국 ESR을 증가시킨다.⁽²⁾

ESR 검사법은 이러한 적혈구 응집도를 관측하여 염증의 정보를 종합적으로 나타낸다. 다양한 질병의 진단 및 추적 관찰에 사용되며, 다른 검사와 병행하여 여전히 널리 사용되는 검사법이다.⁽³⁾

하지만 가장 보편적인 ESR 검사법인 Westergren 방법은 비교적 긴 시간인 1시간이 소요된다. 이에 대한 대안으로 검사 시간이 짧은 TEST 1(Alifax, Padova, Italy)이나 Ves-Matic CUBE 200(Diesse Diagnostica Senese, Siena, Italy)과 같은 자동화된

† School of Mechanical Engineering, Pusan National University (PNU), Associate Professor
E-mail: esyeom@pusan.ac.kr

* School of Mechanical Engineering, PNU, MS student

** Eco-friendly Smart Ship Parts Technology Innovation Center, PNU, Postdoctoral Researcher

*** Department of Oral and Maxillofacial Surgery, Associate Professor

ESR 분석기기 등이 사용된다.^(4,5) 그러나 해당 장치는 광학측정센서 등 특수한 소자를 필요로 하며, 장치마다 일부 영역에서 Westergren 방법으로 얻은 결과와 상관관계가 낮은 구간이 존재한다.

신경망 (Neural network) 모델은 주로 영상 및 음성 신호와 자연어 처리에 사용되며, 난류 모델링 등 유체역학 분야에도 다양하게 적용되고 있다.⁽⁶⁾ 순방향 신경망(Feed-forward neural network; FFNN)과 같은 연산 구조를 가진 신경망은 입력에 대해 과거 정보의 영향을 고려할 수 없지만, 순환 신경망의 종류인 LSTM (Long short-term memory)이나 GRU (Gated recurrent unit)는 순차적인 시퀀스(sequence) 데이터를 처리하는데 우수한 성능을 발휘한다고 보고되었다.⁽⁷⁾

본 연구에서는 실시간 적혈구 침강경향 데이터를 획득하고 GRU 신경망을 사용함으로써 낮은 오차와 진단 시간을 줄인 ESR 예측 모델을 제안한다. 예측 모델의 결과값을 다른 모델과 비교함으로써 제안한 모델의 성능을 평가하고자 한다.

2. 연구 방법

2.1 실험 구성

Fig. 1은 ESR 측정 장치와 영상 획득 과정을 보여준다. 1 mL 용량의 일회용 주사기(Syringe)에 1 mL의 혈액 표본을 주입하여 상온에 수직으로 고정한다. Westergren 방법을 통한 ESR 검사법은 실험 시작 후 1시간 뒤의 침강 거리인 ESR만 측정한다. 본 실험에서는 Westergren 방법과 유사한 실험 장치를 구성하되, 시간에 따른 적혈구의 침강 경향을 확인하기 위해 카메라(D5600, Nikon, Japan)와 컴퓨터를 연결하여 자동화 프로그램(Nikon camera control Pro, Nikon)으로 영상을 획득하였다. 프로그램은 Fig. 1b와 같이 펄스 형태의 트리거(Trigger) 신호를 통한 카메라 제어로 1분 간격으로 60분 동안 60장의 영상을 획득한다. 펄스 신호 위에 위치한 혈액 사진들은 카메라가 시간에 따라 순차적으로 영상을 촬영하는 과정을 보여준다.

130명의 환자로부터 실험용 혈액 표본을 수집하였고 남성 및 여성 비율이 각각 45%, 55%이며

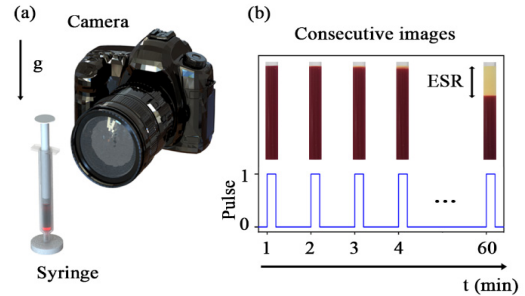


Fig. 1. (a) Schematic setup of ESR measurement. (b) Image acquisition process.

평균 나이가 약 48세이고 표준편차가 약 20이다. 혈액은 채취 즉시 EDTA 튜브에 넣어 항응고 처리되었으며, 실험 직전 원심 분리장치를 통해 혈소판과 백혈구를 분리 및 제거하여 혈장과 적혈구를 헤마토크릿(Hematocrit) 35% 비율로 다시 섞어 실험에 사용하였다. 혈액 표본은 부산대학교병원에서 공급받았으며, 본 연구는 부산대학교병원의 기관 심사위원회(IRB)에 의해 허가되었으며 헬싱키 선언의 규정을 따른다(H-1911-028-085).

2.2 영상 처리 및 데이터 전처리

Fig. 1b에서 확인할 수 있듯이 노란색 혈장과 붉은색인 적혈구의 경계가 침강 경향을 보여 준다. 이를 정량적으로 분석하기 위해 영상처리를 수행하였다. 원본 영상에서 관심 영역을 정하고 각 영상에 Otsu 방법을 사용해 임계값을 지정하여 이진 영상(binary image)으로 변환하였다. 이와 같이 이진 영상에서 중력 방향을 따라 픽셀 값의 변화가 큰 위치를 적혈구와 혈장의 경계로 정의하였다. 이러한 처리를 통해 각 표본의 경계 위치가 60개의 시계열 데이터로 정리된다.

실험 과정에서 모든 표본의 혈액높이를 통일하고, 각 표본의 처음의 경계 위치(최고 높이)를 1로 정규화 하였다. 또한, 신경망의 효율적인 가중치 학습을 위해 최소-최대 정규화 방법을 사용하여 혈액 기증자의 나이 및 성별의 값을 0과 1사이로 조정하였다

2.3 GRU

GRU는 순환신경망의 특성상 데이터에 내재된 선후 관계를 학습할 수 있다. 따라서 적혈구 침강 데이터와 같이 시계열데이터에 적합하다. GRU는 LSTM보다 간결한 구조로서 학습에 필요한 매개변수가 비교적 적지만 유사한 성능을 발휘하는 것으로 실험적으로 입증되었다.⁽⁸⁾ 또한 GRU가 LSTM보다 소규모 데이터셋에 적합한 것으로 보고되었다.⁽⁹⁾

Fig. 2는 GRU 구조를 나타낸다. GRU는 리셋(reset) 게이트(r_t)와 업데이트(update) 게이트(z_t)로 이루어져 있으며 입출력 관계는 다음과 같다.

$$r_t = \sigma(W_r x_t + U_r h_{t-1}) \quad (1)$$

$$z_t = \sigma(W_z x_t + U_z h_{t-1}) \quad (2)$$

$$\tilde{h}_t = \tau(W_h x_t + U_h (r_t * h_{t-1})) \quad (3)$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \quad (4)$$

여기서 리셋 게이트는 직전 시점의 은닉층의 정보를 활용하는 정도를 결정하며, 업데이트 게이트는 식 (2)와 같이 과거와 현재 시점의 정보를 각각 갱신하는 비율을 결정한다. x_t 는 입력 벡터, W 와 U 는 각각 매개변수 행렬과 벡터를 의미한다. σ 는 시그모이드 함수(Sigmoid function), τ 는 하이퍼볼릭 탄젠트 함수(Hyperbolic tangent function), \tilde{h}_t 는 후보 활성화 벡터(candidate activation vector) 그리고 h_t 는 출력 벡터이다.

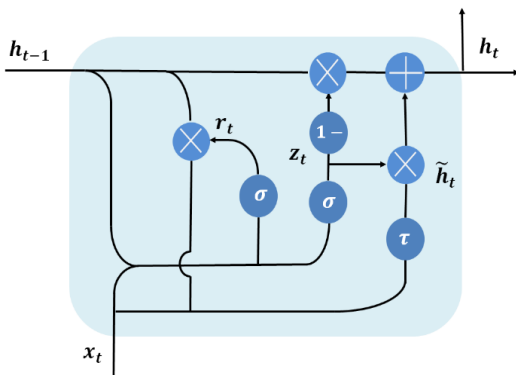


Fig. 2. Structure of gated recurrent unit(GRU) network

2.4 모델 훈련

GRU 모델의 성능을 평가하기 위해 다양한 모델 또한 학습시켰다. 머신러닝 모델에서 우수한 성능을 보이는 랜덤 포레스트(Random Forest)와 신경망 모델로는 순방향신경망(FFNN)과 와이드&딥 신경망(Wide and Deep Neural Network, WDNN)을 비교 모델로서 선택하였다. FFNN은 신경망의 가장 기본적인 모델로서 성능을 평가할 때 기준으로 사용하였다. WDNN은 선형적인 모델과 신경망 모델의 장점을 모두 갖고 있는 것으로 알려져 있어 선택하였다.⁽¹⁰⁾

모델의 손실 함수는 MSE(Mean squared error)를 사용하였으며 아래 식 (5)와 같다. 학습 횟수인 epoch의 최대 크기는 600으로 설정하였다. 그리고 조기 종료(Early stopping)를 통하여 개선이 없는 epoch가 15번 지속될 경우, 학습을 종료하여 연산 부담을 감소하였다. 최적화 기법(Optimizer)은 시행착오법에 의해 ADAM(Adaptive moment estimation)을 사용하였다.

예측 모델의 성능을 평가하기 위한 오차지표는 널리 사용되는 MAE(Mean absolute error)를 사용하였다(식 6). 해당 지표는 실제 값과 예측 값의 오차를 나타내며, 오차 크기의 측도는 절대값이다.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (5)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (6)$$

여기서 n 은 시험 데이터셋의 개수이며, y_i 는 실제 출력 데이터이며, \hat{y}_i 는 모델의 예측값이다.

예측 모델의 학습과 평가를 위해 전체 130개 침강 데이터에서 훈련에 사용된 데이터는 80%인 104개, 시험 데이터셋은 나머지 20%인 26개의 데이터로 나누었다.

3. 결과

3.1 입력 데이터 정의

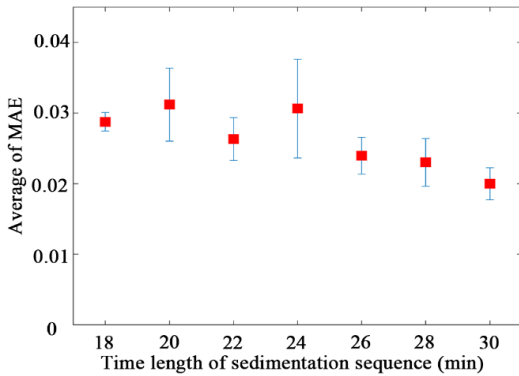


Fig. 3. MAE (mean absolute error) variation of FFNN with respect to length of time

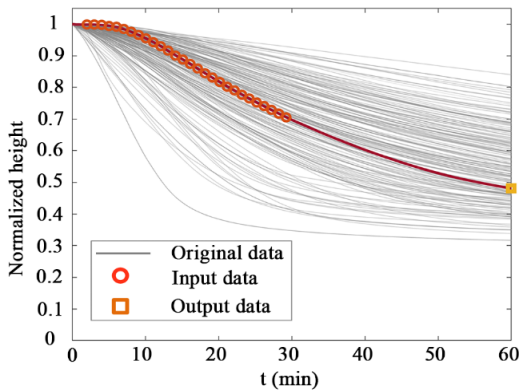


Fig. 4. Temporal variation of sedimentation sequence and resampling

Fig. 3은 ESR 예측 모델에 필요한 시간 길이를 정의하기 위해 MAE로 나타낸 손실과 침강 시퀀스의 시간 길이의 관계를 보여준다. 결과를 얻기 위해 FFNN을 사용하였으며, MAE의 평균과 표준편차는 동일 데이터셋에서 시간 길이를 줄여가며 각각 30번씩 반복하여 계산하였다. 그 결과 입력되는 시간을 줄일수록 MAE의 평균은 상승하며, 분산은 20분과 24분에서 비교적 큰 편이다. 즉 시간을 줄일수록 오차가 상승하는 상충 관계가 존재한다. 본 연구에서는 소규모 데이터셋으로 인한 문제점을 방지하고, 가장 높은 예측 정확도를 위해 시퀀스의 길이는 30분으로 정의하였다.

Fig. 4는 130개의 전체 혈액의 적혈구 침강 시퀀스를 시간 순으로 나타낸 그래프이다. 초기 2분까지의 시간은 제외하고, 모델에 입력되는 시간 길이

는 원형 마커로 표시된 것 같이 2분부터 30분까지로 제한하였다. 출력 값은 사각형 마커로 표시한 ESR(60분 때 침강 길이)이다. 그리고 혈액 표본 기증자의 나이와 성별 또한 모델에 침강 시퀀스 데이터와 함께 입력된다.

3.2 모델 성능 비교

제안한 모델의 성능을 평가하기 위해 랜덤 포레스트(Rand. Forest), FFNN, WDNN 모델과 GRU를 비교하였다. 가중치 학습 시 신경망의 특성인 확률적인 변동을 파악하기 위해 제안한 모델에서 사용된 것과 동일한 입력(시간) 데이터에 대해 각 모델의 전체 학습과정을 40번 반복하고, 시험 데이터에 대한 모델의 예측 결과 또한 40개를 얻었다. 이렇게 얻은 MAE 분포의 평균과 표준 편차를 Fig. 5와 같이 오차 막대로 나타내었으며, 구체적인 수치는 Table 1에 정리하였다. 랜덤 포레스트는 가장 낮은 표준편차를 보여 학습 변동성에 있어서 안정적인 모델이지만, MAE의 평균은 두번째로 커서 성능이 좋다고 판단하기 어렵다. WDNN 또한 MAE의 평균은 랜덤 포레스트나 순환신경망에 비교해서 낮지만, 표준편차가 0.0029로 학습에는 불안정한 모델임을 알 수 있다. 결국 GRU 기반 모델이 평균과 표준편차가 각각 0.0143와 0.0020으로 가장 뛰어난 성능과 비교적 안정된 모델임을 알 수 있다.

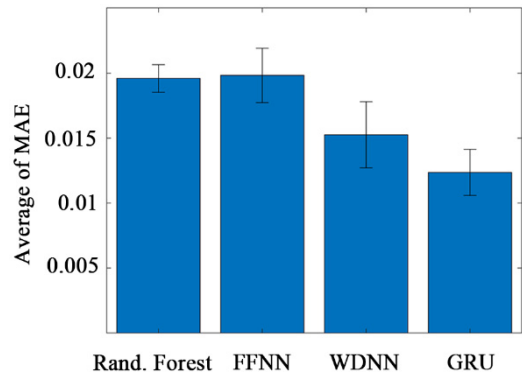


Fig. 5. Statical performance of each model

Table 1. Statistical MAE results of each model

MAE	Model			
	Rand. Forest	FFNN	WDNN	GRU
Average	0.0226	0.0229	0.0176	0.0143
Std	0.0012	0.0024	0.0029	0.0020

4. 결론

본 연구는 염증의 정도를 진단하는 ESR의 예측 모델로 GRU 신경망을 사용하여 진단시간을 1시간에서 30분으로 줄였으며, 제안한 모델의 성능을 다양한 모델과 비교하여 검증하였다. 적혈구의 침강 시퀀스는 시계열 데이터이기 때문에 과거 정보의 영향을 고려하는 GRU 기반 예측 모델이 랜덤 포레스트나 단순 신경망 모델보다 우수한 성능을 보였다. 추후 많은 혈액 표본을 수집하면 진단시간을 지금보다 더 줄일 수 있는 예측 모델을 개발할 것으로 기대한다.

후 기

이 논문은 2019년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(No. NRF-2019R1F1A1062348).

REFERENCE

- 1) Olshaker, J. S. and Jerrard, D. A., 1997, "The erythrocyte sedimentation rate." *The Journal of emergency medicine*, Vol. 15(6), pp. 869-874.
- 2) Stokes, G. G., 1851, "On the effect of the internal friction of fluids on the motion of pendulums."
- 3) Jou, J. M., et al., 2011, "ICSH review of the measurement of the erythrocyte sedimentation rate." *International journal of laboratory hematology*, Vol. 33(2), pp. 125-132.
- 4) Hardeman, M. R., et al., 2010, "Test 1 analyser for determination of ESR. 1. Practical evaluation and comparison with the Westergren technique." *Scandinavian journal of clinical and laboratory investigation*, Vol 70(1), pp. 21-25.
- 5) Curvers, J., et al., 2010, "Evaluation of the Ves-Matic Cube 200 erythrocyte sedimentation method: comparison with Westergren-based methods." *American journal of clinical pathology*, Vol. 134(4), pp. 653-660.
- 6) Brunton, S. L., Noack, B. R. and Koumoutsakos, P., 2020, "Machine learning for fluid mechanics." *Annual Review of Fluid Mechanics*, Vol. 52, pp. 477-508.
- 7) Wu, W., et al. "Using gated recurrent unit network to forecast short-term load considering impact of electricity price." *Energy Procedia*, Vol. 158, pp. 3369-3374.
- 8) Chung, J., et al., 2014, "Empirical evaluation of gated recurrent neural networks on sequence modeling." *arXiv preprint arXiv: 1412.3555*.
- 9) Yang, S., Xueying Y., and Ying Z., 2020, "LSTM and GRU Neural Network Performance Comparison Study: Taking Yelp Review Dataset as an Example." *2020 International Workshop on Electronic Communication and Artificial Intelligence (IWECAI)*. IEEE.
- 10) Cheng, H. T., et al., 2016, "Wide & deep learning for recommender systems." *Proceedings of the 1st workshop on deep learning for recommender systems*, pp. 7-10.