

Deep Learning-Based Detection of Abnormal Respiration Based on Abdominal 3D Shape Dynamically Measured by Moiré Topography

¹Yuki MOCHIZUKI and ^{1,*}Norio TAGAWA

¹ Tokyo Metropolitan University, 6-6 Asahigaoka, Hino, Tokyo, 191-0065, Japan

¹ Tel.: +81-42-585-8416, fax: +81-42-583-5119

* E-mail: tagawa@tmu.ac.jp

Received: 2021 /Accepted: 2021 /Published: 31 October 2021

Abstract: Currently, paramedics cannot make triage judgments in ambulances during emergency transport of infants. If this can be done, it will be possible to select an appropriate destination hospital to provide the most suitable medical care. In this study, we developed a system to detect abnormal respiration by measuring and analyzing three-dimensional abdominal movement using moiré topography and a deep learning model. The performance of the model used in this system was evaluated, and the model was found to be more robust than conventional machine learning models with respect to changes in camera viewpoint.

Keywords: Moiré topography, Active stereo, Deep learning, LSTM, Abdomen shape, Triage, Abnormal respiration.

1. Introduction

Requests for emergency transport of babies and infants are very high, especially at night. In addition, the transportation tends to be concentrated in the major hospitals in the area. Therefore, by appropriately evaluating and diagnosing the condition of them during emergency transportation, they can be transported to a medical institution suitable for their symptoms and provided with necessary medical care. Electronic systems are being developed to continuously monitor the triage vital signs required for them [1]. However, a reliable triage system that can be used adaptively by short-time measurement has not yet been developed. Manual triaging, on the other hand, requires the assessment of respiration, complexion, and consciousness by a skilled healthcare professional. We aim to develop a robust and easy-to-use automatic triage system that can be used in ambulances with poor lighting conditions and high

vibration. As the first step of the research, we investigated a system that measures three-dimensional (3D) movement of the abdomen to detect respiratory abnormalities.

Binocular stereoscopic vision is not sufficient for 3D measurement of the abdomen. That's because the abdomen lacks texture. In recent years, it is conceivable to use an RGB-D camera that has made remarkable progress, but in this study, we use moiré topography. One of the reasons is that in the future, the time variation of the moiré interference fringes before being converted into 3D information may be used for highly sensitive detection of abnormal respiration. In this study, we tried to use a 3D shape reconstructed by moiré topography in order to confirm the intuitive understanding that the time variation of the abdominal shape would be useful for detecting abnormal respiration. In principle, moiré topography cannot determine the absolute 3D shape of an object. To solve this problem and make it easier to measure temporal

changes in abdominal shape, this study proposes a combination of moiré topography and active stereo method.

In this study, a long short-term memory (LSTM) [2] neural network is used to identify abnormal respiration from changes in abdominal 3D shape over time. LSTM models are machine learning models that can make full use of past time series data, which solves the limitation of recurrent neural networks (RNN) that only recent-past data can be used. For the input to the LSTM, the normal vector of the tangent plane at the sampling point on the abdominal surface was used instead of the 3D shape of the abdomen itself. This is because it is taken into consideration that the relative positional relationship between the camera and the target abdomen always fluctuates due to vibration of the environment and changes in the shooting viewpoint during actual use. Through experiments, it was confirmed that the proposed method is more robust to viewpoint changes than the conventional typical machine learning methods.

This paper is an extended version of the proceedings paper [3] at the 7th International Conference on Sensors and Electronic Instrumentation Advances. In particular, the parameter dependence obtained as a result of learning by LSTM is introduced in detail.

2. Measurement System

2.1. Moiré Topography with Active Stereo

In moiré topography, phase values in the range $[0, 2\pi]$ are calculated from moiré interference fringes, and global phase information is restored by unwrapping this local phase information. The depth (distance to the measurement) can be determined by the phase order of each moiré fringe.

The principle of moiré topography (Fig. 1) is that by superimposing a projection of a slit grid pattern on a captured image of the measurement area on which the slit light had been projected, an image with higher order fringes is generated. The lowest order fringe is the moiré fringe.

From the similarity relationship between the triangle ABC and the triangle ASE (Fig. 2), the distance h from the reference grid G to the N^{th} order fringe is calculated by the equation:

$$h_N = \frac{NP_0d}{L - NP_0}, \quad (1)$$

where L is the distance between the light source and the center of the lens, P_0 is the pitch of the reference grid, and d is the projection distance to the reference grid. Since only discrete distances can be measured with Eq. 1, phase information corresponding to the distance between fringes is calculated using a phase-shifting method [4] or the like, such as

$$h(d, \phi, N) = \frac{P_0 d (\phi + 2\pi N)}{2\pi L - P_0 (\phi + 2\pi N)}, \quad (2)$$

where ϕ is the phase between adjacent fringes wrapped from 0 to 2π , and $\phi + 2\pi N$ is the phase measured from the fringes, which is obtained by unwrapping data.

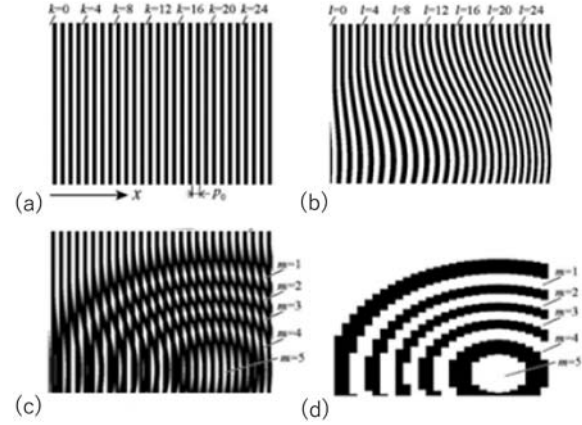


Fig. 1. Generation of moiré interference fringes: (a) projected slit light, (b) captured image with slit light projected onto the target, (c) image with original stripe pattern superimposed on captured image with pattern appearance warped by the targets topography, and (d) extracted moiré interference fringes.

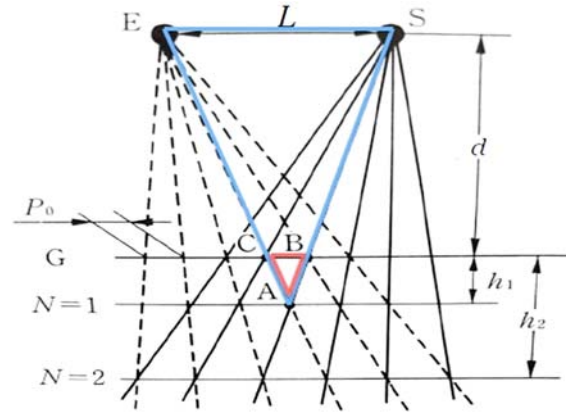


Fig. 2. Principle of distance measurement in moiré topography. S is a point light source, and E is a lens center.

In this study, to obtain local phase information, the sampling moiré method (Fig. 3; [5]) is performed by simple pixel sampling. For each cycle of the captured pattern, N pixels are sampled, and linear interpolation is used to generate N patterns with phase shifts, from which we can calculate the local phase. Fig. 4 shows an example of an abdominal image with captured fringe pattern and the corresponding local phase image. There are several phase unwrapping methods [6-8]. We used the quality map method [6] because of the smooth shape of the abdomen. In this method, reliability of the local phase is calculated (Fig. 5), and the pixels with the highest values are grouped to stack the local phases.

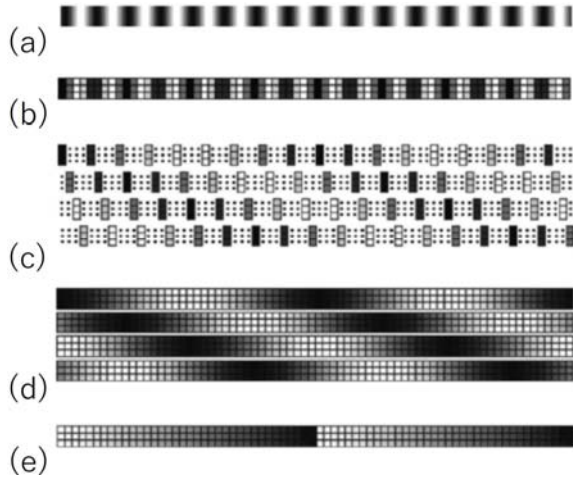


Fig. 3. Phase analysis. (a) original pattern; (b) captured pattern; (c) one cycle of the pattern (N samples); (d) linear interpolation; (e) local phase. (Reprinted from [5]).

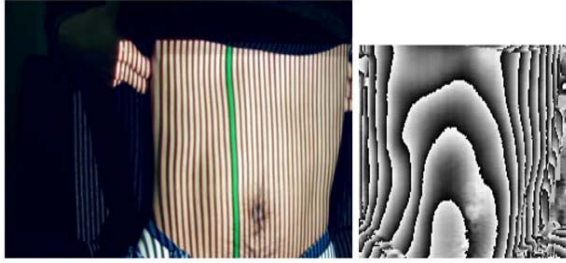


Fig. 4. Captured image (left) and representation of corresponding local phase (right).

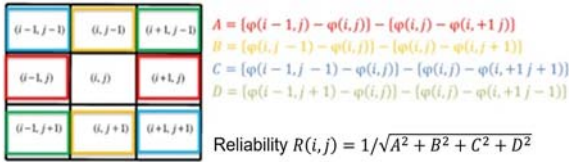


Fig. 5. Reliability definition and the quality map method.

In moiré topography, moiré fringe order cannot be determined, and at least one fringe order, that is, the corresponding distance, must be determined by another method. It is difficult to determine the specific fringe order at each time point because depth measurements, which also change with time, would be necessary. While it is also conceivable to have the patient hold their breath before measurement to determine the order of a particular fringe and then track that fringe in time, this is not realistic. Therefore, we used an active stereo method in which one of the slit lights is colored (green line in Fig. 4), and the depth that colored light is calculated using binocular stereoscopy. This makes it possible to measure depth independently at each time point, which eliminates the need for time-consuming processing. Fig. 6 shows measurement results corresponding to Fig. 4. A measurement error occurs in the slightly shaded area on the right side of the abdominal image in Fig. 4.

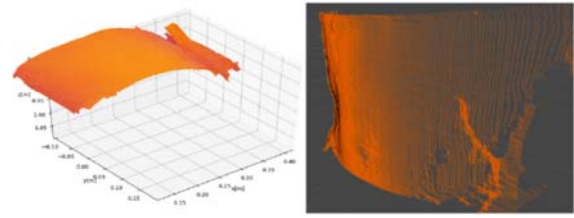


Fig. 6. Resulting 3D measurement of abdominal shape.

However, it is considered that there is no significant effect on the measurement of respiratory status.

A measurement system, with a USB camera and an ultra-small projector controlled by a single-board computer, Raspberry Pi, was manufactured as a prototype (Fig. 7) and used in the experiments described below. The specifications of the camera and the projector used are as follows.



Fig. 7. Moiré measurement system.

Camera

- Model num.: ELP-USBFHD08H-SFV;
- Resolution: 1920 × 1080;
- Sensor: CMOS OV2710;
- Horizontal FOV: 36 deg.

Projector

- Model num.: Ultimems HD301D1;
- Resolution: 1280 × 720, 60 fps;
- Luminous flux: 20 lm ± 10 %;
- Horizontal FOV: 38 deg.

2.2. Detection of Abnormal Respiration

Abnormal respiration such as depressed respiration exhibits different spatial displacements of the abdominal surface than normal respiration—vertical movement of a small area near the epigastrium has an opposite phase to that of the surrounding abdomen during abnormal respiration.

Moiré topography measures the distance from the imaging system to the abdomen. When using this system it is not appropriate to directly use the distance

information for abnormal breathing detection because there are movements of the patient, changes in the imaging viewpoint, and vibrations of the surroundings. The first idea is to convert the distance data into 3D shape data. However, the shape of the abdomen itself is not necessary, and the spatial relationship of temporal unevenness changes is important to monitor the respiratory state. Therefore, in this study, considering the simplification of processing, we decided to calculate and use the unit normal vector on the abdominal surface from the distance data. Namely, time series data consisting of a group of unit normal vectors at multiple points on the abdomen were used as input. Even if the relative positional relationship between the imaging system and the patient's abdomen fluctuates irregularly, the relative relationship of the spatial distribution of the normal vector does not change. Therefore, if such relative relationships are extracted and learned by machine learning without preprocessing such as normalizing the direction of the all normal vectors, less training data is required and an efficient system construction is performed. In the experiments described below, data is collected only when the abdomen and the imaging system are fixed, and machine learning is performed using only that data. Then, the above possibility will be clarified by evaluating the abnormal respiration detection performance in different viewpoints.

Time series data can be handled by normal RNNs. For the periodic data handled in this study, it is important to appropriately extract the necessary information from the past information. Therefore, we use LSTM [2]. An LSTM block (Fig. 8) is used as the intermediate layer of the neural network. The prediction result of the time series is obtained from the previous time series and input, and passed to the next block. Unlike in a normal RNN, in this architecture, a cell state is added to the input element. This cell state stores the LSTM state from the oldest data in the input time series. In the LSTM block, long-term dependence of sequences is learned from the cell state.

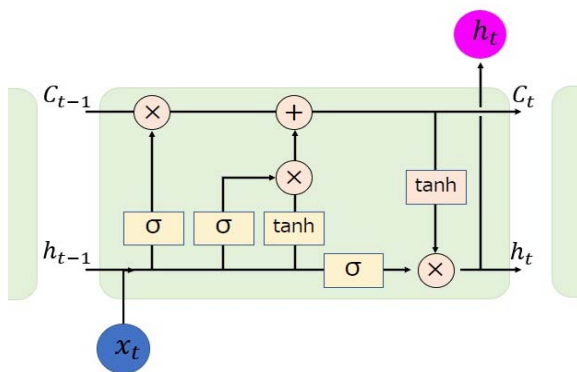


Fig. 8. LSTM block, where h_t is the predicted time series of the input x_t obtained from the prediction result h_{t-1} of the previous time series. A cell state C_{t-1} is added to the input element.

Fig. 9 shows the processing flow of the abnormal respiration detection system. The output in the range $[0, 1]$ corresponds to a probability, where 0 is the probability that respiration is normal and 1 is the probability that respiration is abnormal. The final detection of abnormal respiration will be determined by setting an appropriate threshold for this output. This is beyond the scope of this study and will be the subject of future research. In the experiment, 0.5 was simply set as the threshold.



Fig. 9. Processing flow for abnormal respiration detection based on LSTM.

3. Experiments

3.1. Experimental Conditions

We used normal respiration data to artificially create abnormal respiration data. (Patient participation was not possible as a result of restrictions put in place for the COVID-19 pandemic.) Depressed respiration data were simulated by reversing the vertical movement near the lower part of the sternal pedicle in one cycle of respiration (Fig. 10).

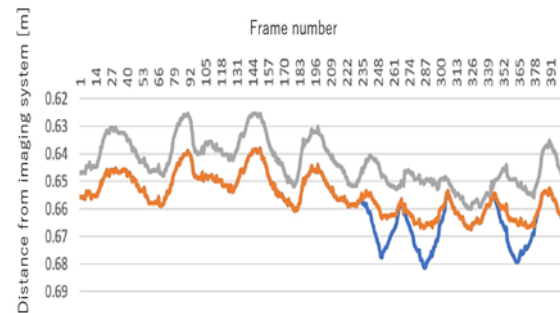


Fig. 10. Time series of distance from the measurement system to abdomen and chest. The gray line is the distance to umbilical region, the orange line is the distance to epigastrium, and the blue is the displacement corresponding to artificially processed depressed respiration.

Examples of 3D abdominal shapes corresponding to normal and depressed respiration are shown in Fig. 11. The vertical movement was in the range (0.5 cm–1.5 cm), and the cycle length was the same as that of the original data.

Measurements were taken from five men (age 20–29 years). For each participant, five measurement sequences of the abdomen and chest were taken, each at a different distance from the camera in 10 cm increments from 50 cm to 90 cm. Each sequence was captured at a rate of 20 frames per second for a duration of 20 seconds. The first

half of the sequence was used as normal respiration; the second half was modified and used as depressed respiration.

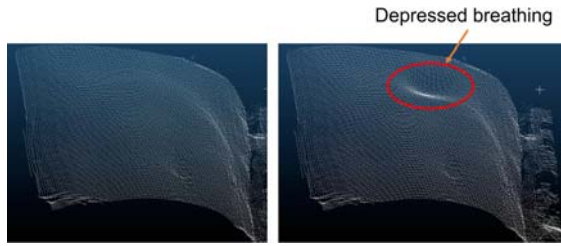


Fig. 11. Examples of abdominal shape for normal (left) respiration and respiratory depression (right). The red ellipse indicates where the phase is opposite to that of the surroundings.

Multiple spatial points on the abdomen, including the epigastrium, were selected by sampling at equal intervals, and the unit normal vector at each sampling point was used as a feature.

Data from four participants (four sequences from each participant's five sequences) were used to train the model, and the remaining sequences (one each from four participants and all five sequences of one person) were used for validation (Table 1). During model training, normal and depressed respiration were input randomly.

Table 1. Details of how to use data for learning and evaluation. A to E indicate the difference between people.

-	50cm	60cm	70cm	80cm	90cm
A	Test	Training	Training	Training	Training
B	Training	Test	Training	Training	Training
C	Training	Training	Test	Training	Training
D	Training	Training	Training	Test	Training
E	Test	Test	Test	Test	Test

The input time series length was 28 frames, the number of sampling points was 32×40 points, the number of hidden layers of LSTM was 384, and the mini batch size was 20 series. The number of pixels constituting the same region (the abdomen extending from the epigastrium to the umbilical region) changed depending on the distance between the measurement system and the target. Since the number of sampling points is predefined, the sampling interval differed for each sequence. Adaptive moment estimation (Adam) was used to optimize learning.

3.2. Experimental Results

First of all, the measurement accuracy of our proposed 3D reconstruction using the manufactured measurement system was evaluated. The measurement accuracy of this system largely depends on the

calibration of the positions and orientations of the projector and camera. After separating the pattern drawn on the calibration jig made of flat plate and the projection pattern generated by the projector, estimate the position and orientation of the projector with respect to the world coordinate system by the method in [9], and then use the method in [10] to estimate the position and the orientation of the camera. Then, they corresponded to each other via the world coordinate system.

In order to confirm the accuracy of the system, the 3D shape of the four-sided pyramid shown in Fig. 12 was measured. The measurement result is shown in Fig. 13.

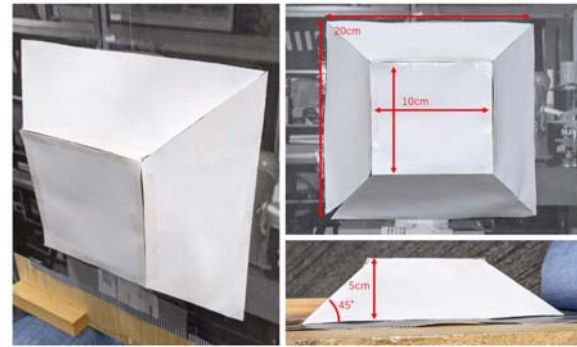


Fig. 12. Measurement target for accuracy verification (four-sided pyramid).

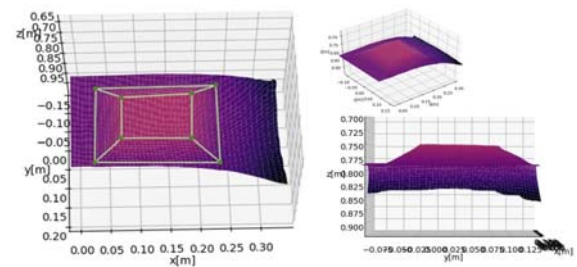


Fig. 13. Measurement result of four-sided pyramid.

From the measurement result, it was found that there was a measurement error of about $+0.5$ cm to $+1.5$ cm in the direction parallel to the depth. It was also found that the error became larger toward the edge. In addition, it was found that there was an error of about -0.95 cm to -0.48 cm in the depth direction. When measuring an object using the active stereo method, which is often used in recent RGB-D cameras, most of the measurement results have an uneven surface, whereas moiré topography using interference fringes obtains a very smooth surface. The above error is larger than the theoretical accuracy of the sampling moiré method. However, due to the characteristics of the proposed method, if an error occurs in the active stereo method, it directly affects the calculation of moiré topography. Since the accuracy of the calibration of the imaging system greatly affects the accuracy of the active stereo

method, it is necessary to confirm the calibration accuracy again. In any case, since the goal of this study is automatic detection of abnormal respiration using time-series data of the abdomen, an error of about 1 cm is sufficiently acceptable.

Next, the experimental results regarding the detection of abnormal respiration will be described. Fig. 14 shows loss convergence during training with the above-mentioned default parameter values. Convergence was fast when the number of layers was large and when the number of sampling points was large; it did not depend on the number of time series.

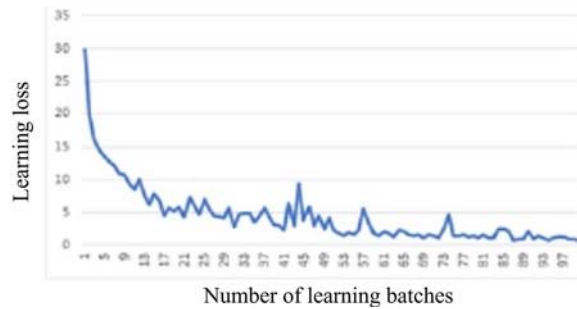


Fig. 14. An example of loss convergence during model training with default parameter values; time series length is 28, the number of sampling points is 32×40 , and the number of hidden layers is 384.

As details, Fig. 15 shows an example of convergence of the loss function during learning when the number of layers is different. As the number of layers increases, the convergence of the loss function becomes slightly faster, but there is a tendency for small increases and decreases. This means that increasing the number of layers, and thus the number of weights in the network, gives the model more freedom to represent and better adapts to the data. According to the discrimination test for the data used for learning, when the number of layers is small, the correct rate with high reliability decreases. However, even if the number of layers is too large, the same correct rate tends to decrease slightly. In this experiment, the correct rate for training data was the highest when the number of layers between 384 and 512. Similarly, we investigated the convergence of the loss function during learning when the number of time series handled by the LSTM was changed. As can be seen from Fig. 16, as the number of time series increases, the overall convergence pattern expands slightly, but there is no significant change. When the number of time series is 20, the loss function fluctuates discontinuously due to the difference in training data. We don't know the exact reason for this, but the time series number 20 may not be appropriate for learning the features common to the data. It was found that the correct rate of normal data decreases when the number of time series is small, although the influence of the number of abnormal data on the correct rate is small. As a result, in this experiment, it was confirmed that

the overall correct rate was high when the number of time series is around 28. Further, Fig. 17 shows the convergence pattern of the loss function when the number of sampling points in the abdomen is increased to 64×80 . It can be seen that the convergence is clearly slower and smoother than that in Fig. 14. The correct rate increased as the number of sampling points increased, but the reliability of discrimination of abnormal respiration decreased significantly. The calculation time is overwhelmingly long when there are many sampling points. Therefore, after that, we set it to 32×40 and conducted an identification experiment on unlearned data.

Performance and processing speed varied depending on the parameters. Since the influence of the time series length was the largest, the results are shown in Table 2 as accuracy. The corresponding inference time is shown in Table 3. The results of the support vector machine (SVM) model, logistic regression (LR), and the hidden Markov model (HMM) are also shown for comparison. The SVM used a second-degree polynomial kernel. LR was optimized using the quasi-Newton method. The HMM was a two-state model. There were no differences in accuracy between LSTM, SVM, and LR models. HMMs, which handle time series information in a manner similar to LSTM models were unexpectedly inaccurate. This may be due to the adoption of a simple two-state model. The longest inference time was for the SVM model, and there was not much difference between LSTM model and LR results. Real-time processing by LSTM is considered possible.

Table 2. Accuracy [%] for different time series lengths. NaN, not a number; LSTM, long short-term memory; support vector machine; LR, logistic regression.

Learning model	Time series length				
	12	20	28	36	NaN
LSTM	91.09	91.76	92.39	92.46	-
SVM	89.26	91.12	92.00	92.46	-
LR	90.45	90.42	91.40	91.91	-
HMM	-	-	-	-	84.41

Table 3. Difference in inference time [seconds] with respect to time series length. LSTM, long short-term memory; CPU, central processing unit; GPU, graphics processing unit; SVM, support vector machine; LR, logistic regression; HMM, hidden Markov model.

Learning model	Time series length				
	12	20	28	36	NaN
LSTM (CPU)	3.1	4.1	6.0	6.7	-
LSTM (GPU)	0.3	0.3	1.4	1.3	-
SVM	428.8	282.2	255.6	238.9	-
LR	1.4	2.2	3.0	4.0	-
HMM	-	-	-	-	30.4

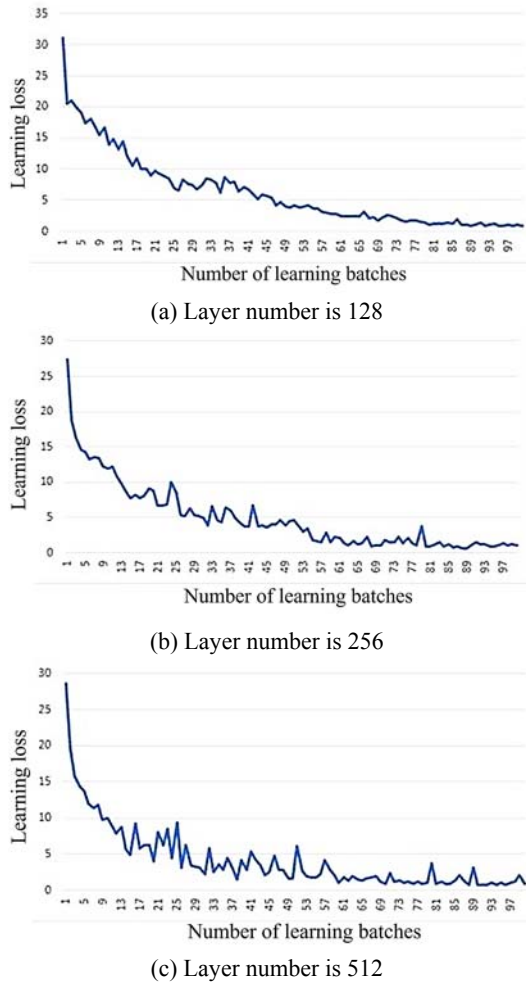


Fig. 15. Changes in convergence of loss function due to the difference in the number of layers.

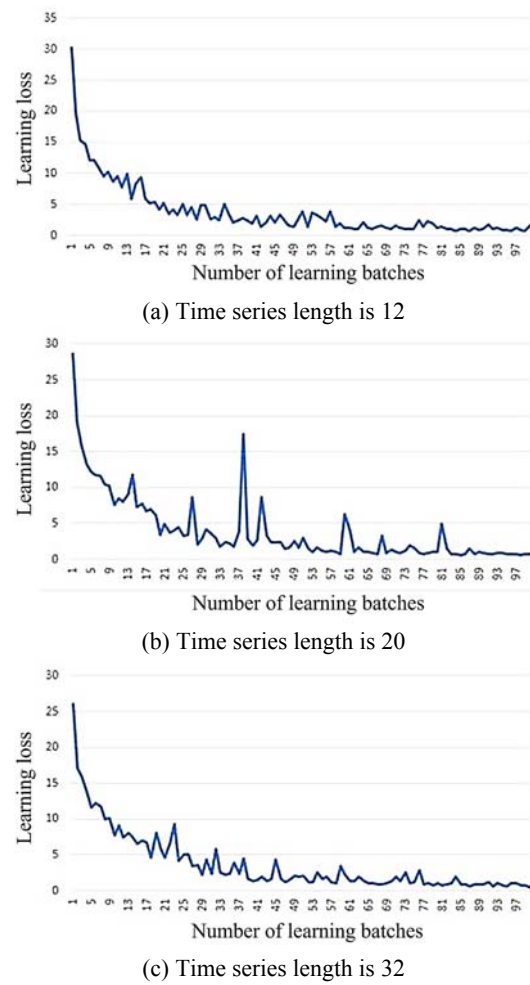


Fig. 16. Changes in convergence of loss function due to difference in the number of layers.

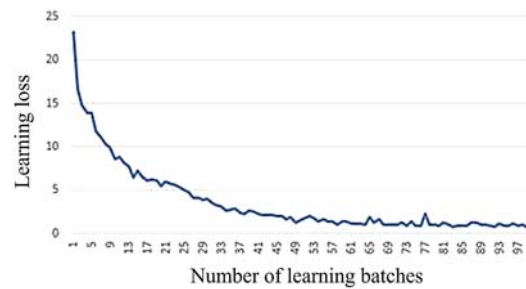


Fig. 17. Convergence of loss function with sampling points of 64×80 .

In actual use, movement of the measurement system and of the patient are expected, and it is difficult to consistently capture images of the abdomen from the front. Therefore, the normal vector calculated when capturing images from the front was rotated to create data corresponding to images captured from an angle. Evaluation data were rotated 30° , -30° , 45° , and -45° around the y-axis to obtain mean accuracy; however, models trained with data measured from the front of the abdomen were used. The LSTM model was able to make robust inferences (Table 4), despite

the change in the camera viewpoint. The HMM was also robust; we consider this to be the effect of the time series information. A strategy for training models using data from multiple directions is also conceivable; this is a topic for the future.

Table 4. Model accuracy. LSTM, long short-term memory; SVM, support vector machine; LR, logistic regression; HMM, hidden Markov model.

LSTM	SVM	LR	HMM
83.15	49.13	47.47	64.55

4. Conclusions

We proposed a method for detecting respiration abnormalities based on 3D measurement of the abdomen using moiré topography and deep learning. The active stereo method was used additionally for moiré topography to determine distance to the reference plane and the relative fringe order. Many other methods for abdominal shape measurement have been proposed [11-15]. Our method of detecting

abnormal respiration is robust to changes in camera viewpoint.

An inexpensive RGB-D camera does not differ much in price from the moiré measurement prototype developed in this study. However, our measurement system can reduce the computational load by using moiré fringes directly to detect abnormal breathing without reconstructing the 3D shape of the abdomen. To this end, it is important to deal with dynamic changes in the relative positional relationship between the imaging system and the abdomen. In addition, dynamic changes in moiré fringes are expected to be more sensitive to abnormal respiratory conditions than changes in the 3D shape of the abdomen. This is an interesting future technical issue.

In this study, the measurement system was used at 20 fps, so the length of the time-series data evaluated for use in LSTM was around 1 second. This time corresponds to about one breathing cycle, as breathing is fast when suffering from an illness. Therefore, in this experiment, abnormal respiration is determined from one respiratory state, and features such as the volatility of the respiratory cycle are not used. In future studies, we will explore the possibility of detecting long-term features of abnormal respiratory conditions, such as changes in respiratory periodicity and the frequency of depressed respiration.

The realization of automated triaging judgments with machine learning is a complex task. For example, triage judgments are also made by integrating other information such as complexion. We used measurement of depressive breathing as the first step in developing an automated triage system.

Acknowledgements

We would like to express our deep gratitude to Dr. Masataka Honda, Professor of Tokyo Metropolitan Children's Medical Center and Dr. Naoki Shimizu, Professor of St. Marianna University School of Medicine, for their suggestions and guidance on this research subject.

References

- [1]. K. Sakanushi, T. Hieda, T. Shiraishi, Y. Ode, Y. Takeuchi, M. Imai, T. Higashino, H. Tanaka, Electronic triage system for continuously monitoring casualties at disaster scenes, *Journal of Ambient Intelligence and Humanized Computing*, Vol. 4, Issue 5, 2013, pp. 547-558.
- [2]. K. Greff, R. K. Srivastava, J. Koutnik, B. R. Steunebrink, J. Schmidhuber, LSTM: a search space odyssey, *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 28, Issue 10, 2017, pp. 2222-2232.
- [3]. Y. Mochizuki, N. Tagawa, 3-D measurement of abdomen by moiré analysis and its use for abnormal respiratory conditions by deep learning, in *Proceedings of the 7th Int. Conf. Sensors Engineering and Electronics Instrumentation Advances*, Palma de Mallorca, Spain, 22-24 September 2021, pp. 83-87.
- [4]. L. Jin, Y. Kodera, T. Yoshizawa, Y. Otani, Shadow moiré profilometry using the phase-shifting method, *Optical Engineering*, Vol. 39, Issue 8, 2000, pp. 2119-2123.
- [5]. S. Ri, M. Fujigaki, Y. Morimoto, Sampling moiré method for accurate small deformation distribution measurement, *Experimental Mechanics*, Vol. 50, Issue 4, 2010, pp. 501-508.
- [6]. M. A. Herraiez, D. R. Burton, M. J. Lalor, M. A. Gdeisat, Fast two-dimensional phase unwrapping algorithm based on sorting by reliability following a noncontinuous path, *Applied Optics*, Vol. 41, Issue 35, 2002, pp. 7437-7444.
- [7]. K. Sartor, G. B. Tenali, A. Peter, J. D. V. Allen, M. Rahmes, IFSAR processing using variational calculus, in *Proceedings of the American Society for Photogrammetry and Remote Sensing Annual Conference*, Portland, 28 April-2 May 2008.
- [8]. L. F. Wang, M. Yan, Weighted Kalman filter phase unwrapping algorithm based on the phase derivative variance map, *Applied Mechanics and Materials*, Vol. 475-476, 2014, pp. 991-995.
- [9]. Z. Lei, X. Liu, W. Lu, Z. Lei, L. Chen, L. Zhou, A new calibration method of the projector in structured light measurement technology, *Procedia CIRP*, Vol. 27, 2015, pp. 303-308.
- [10]. Z. Zhang, A flexible new technique for camera calibration, *IEEE Trans. PAMI*, Vol. 22, Issue 11, 2000, pp. 1330-1334.
- [11]. K. Povsic, M. Jezersek, J. Mozina, Real-time 3D visualization of the thoraco-abdominal surface during breathing with body movement and deformation extraction, *Physiological Measurement*, Vol. 36, 2015, pp. 1497-1516.
- [12]. V. D. J. Sam, J. J. J. Dirckx, Real-time structure light profilometry: a review, *Optics and Lasers in Engineering*, Vol. 87, 2016, pp. 18-31.
- [13]. L. Lu, Y. Ding, Y. Luan, Y. Yin, Q. Liu, J. Xi, Automated approach for the surface profile measurement of moving objects based on PSF, *Optics Express*, Vol. 25, Issue 25, 2017, pp. 32120-32131.
- [14]. M. Zhong, X. Su, W. Chen, Z. You, M. Lu, H. Jing, Modulation measuring profilometry with auto-synchronous phase shifting and vertical scanning, *Optics Express*, Vol. 22, Issue 26, 2014, pp. 31620-31634.
- [15]. H. Wu, S. Yu, X. Yu, 3D measurement of human chest and abdomen surface based on 3D Fourier transform and time phase unwrapping, *Sensors*, Vol. 20, Issue 4, 2020, pp. 1091.

