# Action Primitive Classification of Sewing Operations toward Training Unskilled Workers

*Akihisa Ishikawa, Naoki Hiratsuka*
*Graduate School of Science and Technology*

*Yutaka Takase, Solvi Arnold, Kimitoshi Yamazaki*
*Faculty of Engineering*

*Shinshu University*
*Wakasato 4-17-1, Nagano, Nagano, Japan*
kyamazaki@shinshu-u.ac.jp

***Abstract* – In this study, we proposed a system for measuring hand posture and fingertip force for training unskilled workers who sew with sewing machines. Furthermore, we classified sewing behaviors by constructing deep learning models using time-series color image data, hand posture, and fingertip force as inputs. Our experiments demonstrated that the classification model based on color images is the highest accuracy. Meanwhile, using the fingertip force-based model can not only classify sewing actions but can also extract the operator's motion data in detail.**

***Index Terms – sewing operation, fingertip force sensing, action classification.***

## I. INTRODUCTION

As technology advances, manufacturing operations are becoming increasingly automated. Conversely, there are still many situations that are difficult to automate and require human skills. An example of such operation is the sewing operation of cloth and other materials. In the manufacturing of fabric products, flexible objects are manipulated using various fingering techniques. It can be imagined that the judgment of action is based not only on visual information but also on forces applied to the fingertips. In other words, to convert a person's work ability into data and effectively communicate it to others, it is necessary to perform multimodal measurements, analyze the results, and present them in an appropriate format.

We are interested in developing a training system that can help unskilled workers efficiently acquire the ability to operate flexible objects such as those used in sewing. The purpose of this study is to establish a motion classification method, which is an important element for achieving this goal. Motion classification here means that when time-series data of a worker's observations are input at regular intervals, a class number can be obtained to indicate what the worker is currently doing. In other words, it is possible to accurately determine what action was being taken at what time. In the future, we will develop a training system for unskilled workers based on this method. We hope to contribute to alleviating the shortage of workers in this industry.

Studies have been conducted on measuring and analyzing human behavior for skill extraction. However, most of these studies have focused on sports or manipulation tasks for rigid body objects [1,2]. Conversely, we target the sewing task. In this case, operators are required to sew together flexible fabrics using a sewing machine. Therefore, a worker's body movements are fundamentally small, and the work proceeds through fine movements of the fingers. It is a new challenge to perform appropriate motion classification under these conditions.

In addition, sewing work entails the know-how to properly handle a deformable object such as cloth. It is desirable to present data that enable non-skilled workers to acquire such know-how when they receive training. In this study, we do not go so far as to show the know-how, but we target the measurement of hand posture and the force applied to the fingertips. This approach allows for the presentation of data on subjects that cannot be expressed by simply presenting the results of motion classification.

The contributions of this study can be summarized as follows.

- We propose a method for classifying worker behavior using time-series data measuring sewing operations as input.
- The data used as cues for classification are image sequence, hand posture, and fingertip force. We developed a classifier that considers each characteristic.
- Data on actual sewing operations were collected to verify the classification accuracy of the proposed method. We also visualized some results to confirm the possibility of using the data.

This paper is organized as follows. The next section describes related work. Section III describes the problem setting and our approach. Section IV details the classification method based on image sequences, and Section V details the classification method based on hand posture and force data. Section VI reports the actual data measurement and classification results, and Section VII summarizes the study.

## II. RELATED WORK

Action classification has been attempted for tasks performed at hand, such as factory assembling work. Malc et al. [3] classified the type of work in an assembly task with screws and bolts using image-based part recognition and an inertial measurement unit (IMU) sensor attached to a human wrist. Chengjun et al. [4] recognized assembly tasks with repetitive motions based on image-based tool recognition and human skeleton estimation. These studies recognized actions by observing both the human and the manipulated object but assumed that the shape of the manipulated object is not deformed. However, in our study, it is necessary to be able to recognize actions even if the manipulated object is deformed because it is assumed that flexible objects such as cloth are being handled.

Azadi et al. [5] used an IMU sensor attached to a worker's wrist to detect frequent small repetitive movements and cluster them, such as screwdrivers and wrenches. Riedel et al. [6] recognized assembly tasks by recognizing the shape of the worker's hands using a camera and classifying them into five predefined basic movements. However, these studies focused only on the operator's movements and did not consider the force applied to the hands and fingertips. However, in flexible object manipulation, it is necessary to measure the force applied to the fingertips because the shape of the flexible object changes significantly depending on the magnitude of the applied force.

Conversely, there have been attempts to classify motions by considering the force applied by the worker. Mo et al. [7] measured the hand posture and fingertip force of a worker performing a sewing operation using a minimally invasive measurement method that does not impair the operator's finger sensation. The effectiveness of their method was verified in a simulated sewing operation by segmenting a series of operations. Fermüller et al. [8] demonstrated that capturing images of a worker's hand and simultaneously measuring the force applied to the fingers improves the accuracy of action prediction. Becker et al. [9] showed that a band composed of myoelectric sensors attached to a person's arm can estimate the finger movement and the force applied to the finger and demonstrated various applications, mainly tablet terminal operations. In this way, the estimation of fingertip force has shown the potential to benefit work process recognition and subsequent application, even if the rigidity of the manipulated object is low. However, because these studies did not focus on actual flexible object manipulation tasks, they did not show the actual operator's movements and force applied to the fingertips.

## III. Issues and Approach

### A. Issues in Action Primitive Classification

This study focuses on sewing operations. In the sewing task, fabric parts are sent to a sewing machine and sewn with threads. We aim to record this workflow as appropriate data to classify the primitive actions. Furthermore, it is desirable to measure the worker's actions and estimate the status of each action. One of the most promising clues for this purpose is the movement of the fingers. In particular, the tips of the fingers have an important role in the operation of a sewing machine. It would be good to know in what posture the fingers are positioned to realize operations with the fingertips. It would also be useful to know in which direction the force is applied to the fingertips during operation. The issues are what kind of data should be used for classification, and how to measure and utilize finger posture and fingertip force.

Moreover, one way to determine finger posture is to use a wearable measurement device such as a data glove. However, sewing requires fine finger movements; thus, wearing data gloves may interfere with the original movements of the fingers. Furthermore, the measurement of the force applied to the fingertips must be devised. Sewing work often relies on the sensation of the fingertips touching the fabric. Therefore, it is desirable to adopt a method for measuring the applied force without losing the fingertip sensation.

### B. Our Approach

Based on the discussion in the previous subsection, the measurement approach for this study is as follows.

1. An RGBD camera is placed in a position which both the operator's hand and the sewing machine are visible, and color and depth images are acquired in time series. However, it is unavoidable that a part of the hand may be hidden depending on the work situation.
2. The hand posture is estimated using RGB images, and then, the results are converted to three-dimensional (3D) information using depth images.
3. For the force applied to the fingertips, we adopt the method of attaching strain sensors to the fingernails [13]. This method enables the measurement of the force applied by the fingertip without losing the sensation of the finger pad.

Next, we describe the action primitive classification approach. The data obtained as described above are used as input to classify the action elements. Classes are manually defined on the basis of features of the fingertips and hand movements. The classifier is constructed so that the current action class is output when time-series data are input at regular intervals.

The input data for the classifier are a sequence of color images, hand posture, and fingertip forces. The sequence of color images records not only the operator's movements but also the operation of the sewing machine and fabric parts. The image sequence is also expected to provide an appropriate output when there are frames for which the hand posture cannot be estimated or the quality of the force data is unsuitable for action classification. Conversely, hand posture and fingertip forces are useful for representing the quality of the task in detail. Thus, they can be used for classification, or to visualize the skill of the operator appropriately. We discuss the classifiers and visualization of each data, believing that utilizing the advantages of each data will lead to the construction of a better training system.

## IV. Image-Based Classification

### A. Class Definition

In sewing operations using general household sewing machines, there are multiple actions even when sewing a single piece of cloth. In such a case, it is desirable to be able to classify which process the worker is in and what the worker is operating at a certain time. Therefore, in this study, we define seven kinds of action primitive for sewing work (Fig. 1).

1. Prepare fabric parts to be sewn and place them on the sewing machine,
2. Lift/drop the presser foot,
3. Send fabric parts to the sewing machine for sewing,
4. Rotate the cloth parts with the needle stuck in them,
5. Make backstitching,
6. Cut yarn tails,
7. Press the sewing start/stop button.

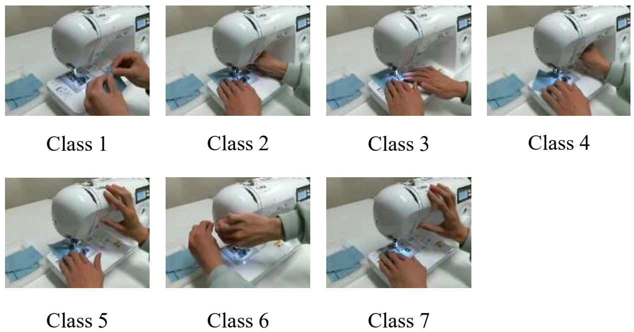



Fig. 1 Class definition of sewing task

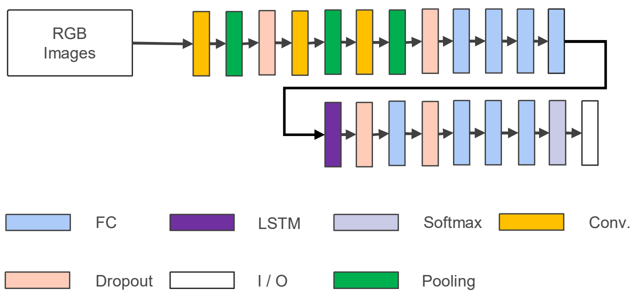



Fig. 2 Image based classification model

### B. *Classification Method*

Figure 2 shows the structure of the image-based classification model used in this study. The proposed classification model is a deep learning model that uses time-series images as input. The input image sequence consists of multiple images taken within a certain time span, and the order in which the images were taken is maintained.

First, a convolution operation is applied three times to each image in the input image sequence in a convolutional layer. Next, the image is compressed to 64 dimensions using a fully connected layer, and then input to a long short-term memory (LSTM) layer. Then, a fully connected layer is applied several times for classification. The output is seven-dimensional. Thus, after applying the convolutional and pooling layers, the LSTM layer is applied to extract spatial feature from each of the images and temporal feature from the image sequence.

The input to the classification model is a set of images compressed to 120 × 85 dimensions and consisting of an image sequence of approximately 0.2 s in length. The use of compressed images reduces the weight of the classification model and speeds up the inference.

### C. *Classifier Training*

The sewing task setup in this study requires several tens of seconds per trial. However, the time required for each class varies. As a result, the amount of data in each class is imbalanced. Because training on imbalanced data significantly affects the performance of the classification model, this study uses undersampling to randomly remove data from the class with the largest amount of data.

The cross entropy loss is used as the loss function for deep learning. The reason is that the cross entropy loss, which is implemented in PyTorch [10], the deep learning library we use, is synonymous with the inclusion of a softmax function and is therefore generally easy to use in classification problems. In addition, Adam [11] is used an optimization method.

The training termination judgment is limited to 100 epochs, and the model with the lowest loss is adopted as the classification model.

## V. Measurement and Classification Based on Hand Posture And Fingertip Forces

### A. *Hand Posture Estimation*

To understand the motion of the hand in detail, we would like to know the 3D motion of the worker's hand. In particular, the positional relationship between the finger joints is important. Therefore, we estimate a skeletal model using RGB images obtained from a 3D range image sensor as an input. Although many methods have been proposed for estimating skeletal models, we adopt MediaPipe Hands [12] because of its fast-processing speed, ease of execution, and high estimation accuracy. 21 landmarks are obtained per hand with MediaPipe Hands. In this study, we use the two-dimensional image coordinates of the estimated landmarks and the corresponding depth image as input and apply a pinhole model to obtain the 3D positions of the landmarks as seen from the camera coordinate system.

However, the handedness labels output by MediaPipe Hands are typically inaccurate. Therefore, we attached different augmented reality (AR) markers to the left and right wrists and identified the left and right hands by determining the distance between the AR markers and each wrist landmark output by MediaPipe Hands.

### B. *Writhe Matrix*

There are individual differences in the size of a worker's hands. In addition, it is difficult to perfectly match conditions such as camera position and angle. Therefore, it is desirable to develop a hand posture expression method that minimizes the effects of individual differences and camera viewpoints.

Vinayavekhin et al. [14] successfully reproduced the motion of a human finger in a simulator with a real robot hand that has a different joint structure by using the Writhe Matrix [15] to represent the posture relationship between a linear object and a human finger. The Writhe Matrix used in their study represents the relative relationship between linear objects and is unaffected by the camera viewpoint, making it an effective method for representing hand posture in our study.

In this study, we represent hand posture using Writhe Matrix. First, two fingers $h^A$ and $h^B$ are selected, and the line from the wrist landmark to each fingertip landmark are, respectively, regarded as single lines. Next, by considering the $m$-th skeleton of the selected fingers as a line segment $h_m$ and the $n$-th

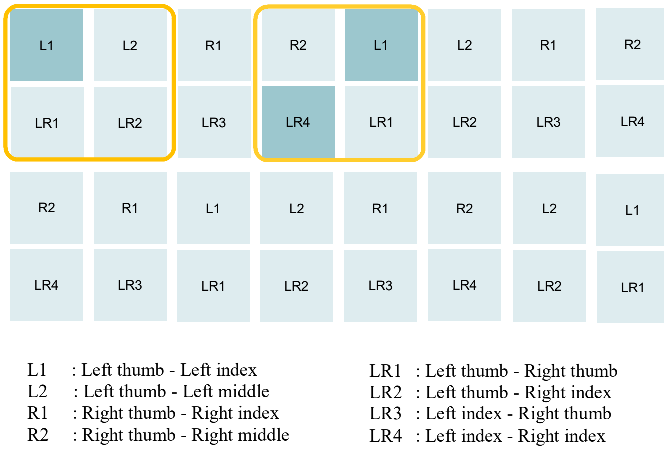



Fig. 3 Writhe Matrix and CNN kernel

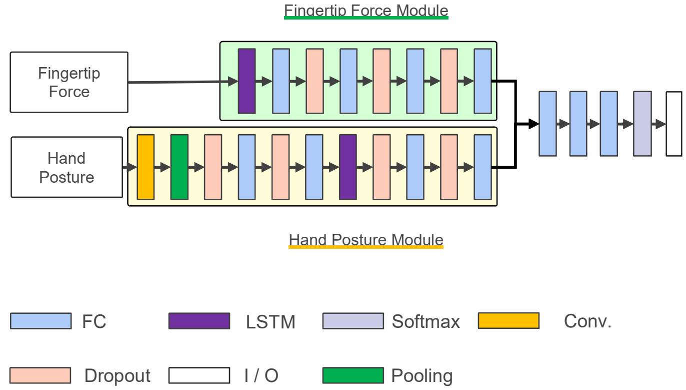



Fig. 4 Classification model of hands posture and fingertip force

skeleton of the selected fingers as line segment $h_n$, the Gauss linking integral (GLI) approximation $g_{mn}$ is calculated for both skeleton combinations. These GLI approximations $g_{mn}$ are arranged in a table to construct the Writhe Matrix for fingers $h^A$ and $h^B$. The resulting Writhe Matrix is a 4 × 4 matrix. In the same way, Writhe Matrices are created for other finger combinations.

The constructed Writhe Matrices can be arranged in the depth direction and handled as a multi-channel image. In this study, however, eight Writhe Matrices constructed by combing the left and right thumb, index fingers, and middle fingers are aligned vertically and horizontally. Then, the relationships between the Writhe Matrices are extracted using a convolutional neural network (CNN). Furthermore, by repeatedly arranging the same Writhe Matrices, as depicted in Fig. 3, distant Writhe Matrices, such as L1 and LR4 in Fig. 3, are placed close to each other and fit into an 8 × 8 kernel.

### *C. Fingertip Force Estimation*

It is necessary to estimate the force applied to the fingertips without losing the operator's finger sensation. To achieve this, we use nail-attached fingertip force sensors developed by Yamazaki et al [13]. This sensor measures the strain of a fingernail by attaching two strain gauges to it and estimates the force applied to the fingertip.

A low-pass filter is applied to the obtained strains to pass only frequencies below 25 Hz. Then, the transformation parameters obtained from the calibration are used to estimate the fingertip force. The method of obtaining the transformation parameters in the calibration is as follows. First, a finger with a nail-attached contact force sensor is pressed down vertically several times on a three-axis tactile sensor placed on a horizontal board. There is a logarithmic function relationship between the amount of strain $y$ and the fingertip force $x$ [13]. As a result, $a$ and $b$ in the following equation are determined using the least-squares method,

$$y = a \log x + b. \tag{1}$$

From then on, the experimentally measured strains are converted to fingertip forces using $a$ and $b$.

### *D. Classification Method*

Figure 4 depicts the structure of the hand posture and a fingertip force-based classification model. This model uses time-series data of hand posture and estimated force as inputs. First, we describe the hand posture module shown in Fig. 4. The 16 × 32 dimensional matrices constructed in Section V-B are input to a convolutional layer. Each matrix is reduced to 64 dimensions by fully connected layers. After that, the matrices are input to an LSTM layer to extract time-series information. The output of the hand posture module is 32 dimensions.

Next, the fingertip force module shown in Fig. 4 is described. The fingertip forces estimated from the amount of nail strain are input for some time. Then, an LSTM layer and fully connected layers are applied in order. The output of the fingertip force module has 32 dimensions.

Finally, the hand posture and fingertip force module are combined and passed through several fully connected layers to create a classifier. The input has 64 dimensions, and the output has seven dimensions.

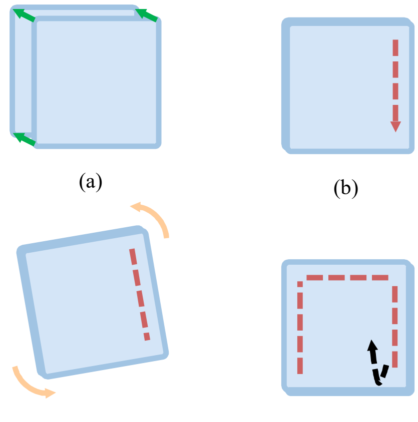
## VI. Experiments

### *A. Sewing Tasks and Measurement Settings*

To verify the effectiveness of the proposed method, we establish a sewing task using a household sewing machine. In this task, two rectangular fabric parts are stacked on top of each other and sewn on three sides. Figure 5 shows the sewing procedure.

First, two pieces of fabrics were stacked, as depicted in Fig. 5 (a). Next, the pair of fabrics was sewn along the dashed lines (Fig. 5 (b)) and were turned (Fig. 5 (c)). These processes were repeated twice. After the other side was sewn, a backstitch was made (Fig. 5 (d)), and the threads were cleaned up at the end.

Next, the measurement environment for image-based classification is described. The image sensor is placed on the left side of the operator so that the various buttons, sewing



(a) (b)

(c) (d)

Fig. 5 Sewing task procedure

(a) (b)

Fig. 6 Camera perspective of each method

needles, fabric parts, and the operator's hand can be seen. It is unavoidable that the presser foot lever is not shown in the image. Figure 6 (a) depicts the image obtained from the placed image sensor. The image sensor is a Logitech C615n, which operates at 30 fps to acquire color images.

The measurement environment for classification based on hand posture and fingertip force is described. First, the worker's hand posture was acquired using Azure Kinect, a 3D range image sensor installed at approximately 0.2 m above the worker's head at an angle of $-60°$. This arrangement was designed to measure the hands without interfering with the operator's work. The 3D range image sensor was used to obtain pairs of color images and depth images at a frame rate of 10 fps. Figure 6 (b) depicts the color image obtained from the placed 3D range image sensor. The nail-attached fingertip force sensors were attached to the middle and index fingers and thumb of the right hand and the index finger of the left hand. The strain gauge used was KFGS-3-120-C1-27, manufactured by Kyowa Electric Co., Ltd. The NR-ST04 strain measurement unit manufactured by KEYENCE Co., Ltd., was used as data logger for the strain gauges, and measurements were performed at 1000 Hz. The estimated fingertip forces were taken from 0.5-s time-series data at regular intervals, and 64 values were considered one sequence of data.

## B. *Data Collection and Learning*

Table I Classification result

| | recall | precision | f1 score |
|---|---|---|---|
| image based | 0.940 | 0.940 | 0.940 |
| posture based | 0.432 | 0.382 | 0.389 |
| force based | 0.628 | 0.641 | 0.624 |
| concatenate | 0.665 | 0.726 | 0.687 |

To evaluate the effectiveness of the classification method based on images, we measured 20 sewing tasks, from three subjects (male, 20's), each of whom completed 20 trials. However, we excluded 30 trials in which the intended sewing task could not be completed because of procedural errors or equipment malfunctions. As a result, the number of usable training data points was 30 trials, which were converted to 17,689 sequence data points. In addition, we measured one sewing task from each subject on a different day as test data.

To investigate a classification method based on hand posture and fingertip force, we measured 120 sewing tasks from four subjects (male, 20's), each of whom completed 30 trials. However, 29 trials were excluded in which the intended sewing task could not be completed because of procedural errors. In addition, two trials for each subject were removed as test data for the final evaluation. Therefore, the total usable training data points was 83 trials. If either the hand posture or strain gauge output was missing for more than half of the trials in a sequence, it was excluded from the dataset. As a result, the number of available training data pints in terms of sequence data was 136,538.

Next, we describe the learning method for the classification method based on hand posture and fingertip force. In constructing the training and test data, the sequence data were cut out according to the number of fingertip force data points with the highest acquisition frequency. This is to assume that the latest data are always input when the classification model is used in actual factories.

## C. *Classification Results*

Table I shows the classification results for each model when the test data were input. For methods other than the image-based model, cross-validation was performed by dividing the training data into five parts, and the average of the evaluation values is shown. According to the results, the image-based method had an F1-score of 94%, indicating high performance. The classification method based on hand posture and fingertip force had an F1-score of 68.7%. The hand posture-based model had the lowest F1-score at 38.9%. Further investigation is required to determine whether this is due to measurement errors caused by occlusion.

In the fingertip force module, the confusion between class 2 (lift/drop presser foot) and class 7 (press a button) is noticeable. This may be because both tasks were similar in that a strong force was applied to the fingertips of the thumb. However,

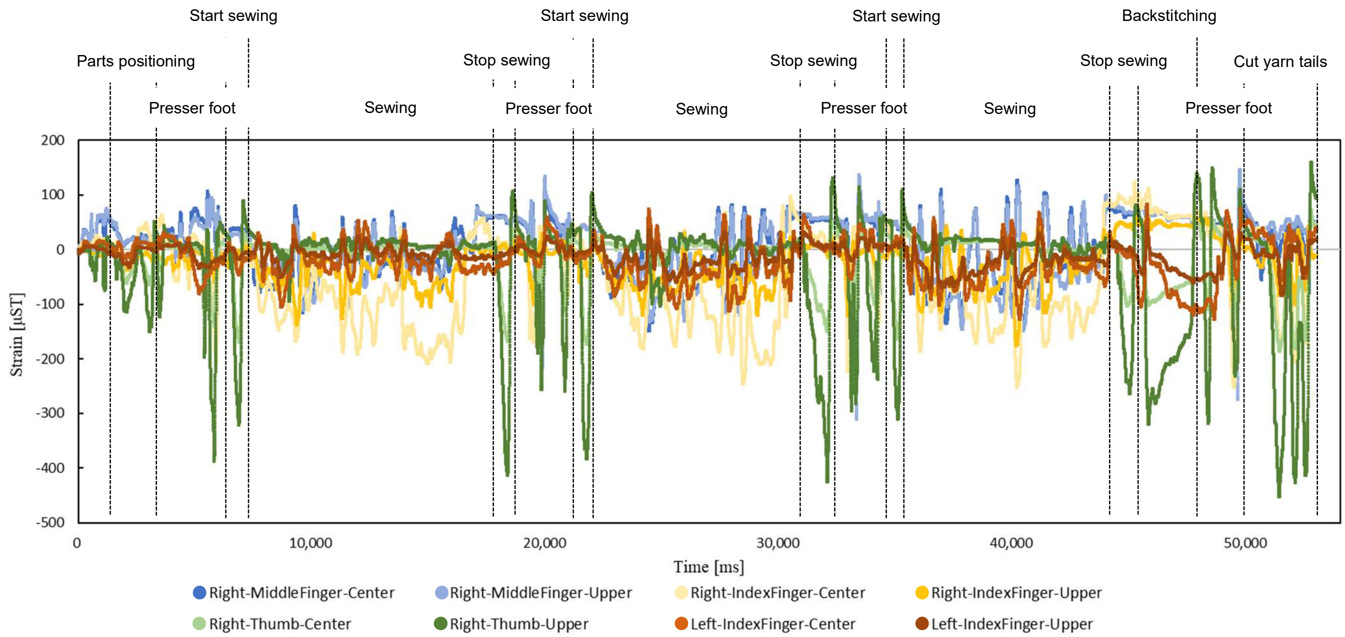



Fig. 7 Measurement of sewing work

Table II Confusion matrix of fingertip force model

| | class 1 | class 2 | class 3 | class 4 | class 5 | class 6 | class 7 |
|---|---|---|---|---|---|---|---|
| class 1 | 6895 | 68 | 0 | 0 | 0 | 0 | 0 |
| class 2 | 701 | 5409 | 368 | 714 | 485 | 377 | 1281 |
| class 3 | 246 | 2235 | 26402 | 498 | 152 | 0 | 2417 |
| class 4 | 0 | 954 | 36 | 745 | 21 | 0 | 524 |
| class 5 | 0 | 272 | 41 | 16 | 3667 | 358 | 695 |
| class 6 | 0 | 261 | 105 | 58 | 467 | 2568 | 85 |
| class 7 | 349 | 1365 | 1750 | 218 | 401 | 17 | 7452 |

Table III Confusion matrix of concatenate model

| | class 1 | class 2 | class 3 | class 4 | class 5 | class 6 | class 7 |
|---|---|---|---|---|---|---|---|
| class 1 | 7333 | 196 | 0 | 0 | 0 | 0 | 0 |
| class 2 | 329 | 6495 | 57 | 1230 | 247 | 442 | 571 |
| class 3 | 343 | 1478 | 27017 | 468 | 182 | 477 | 1980 |
| class 4 | 0 | 868 | 36 | 1166 | 8 | 0 | 182 |
| class 5 | 0 | 143 | 139 | 38 | 3693 | 479 | 536 |
| class 6 | 0 | 109 | 233 | 63 | 325 | 2213 | 19 |
| class 7 | 298 | 811 | 887 | 422 | 1688 | 107 | 7337 |

according to the confusion matrix of the combined model shown in Table III, the confusion between class 2 and 7, which had low classification accuracy in the fingertip force module alone (Table II), has been improved.

*D. Action Data Analysis*

This section discusses the movement of the fingers and the force applied to the fingertips during the sewing task. Figure 7 depicts an example of the strain gauge output obtained from sensors attached to each finger during a sewing task. For example, at around 2,000 ms, the strain of the right thumb, right index finger, and left index finger changes in the area where the fabric parts were stacked. This is because the positions of the clothes were adjusted so that the fingertips were pressed together, as depicted in Fig. 8 (a).

During the first manipulation presser foot (approximately 6,000 ms), the strain output of the right thumb changed significantly, and the strain of the left index finger also changed simultaneously. That is because the right thumb was used to drop the presser foot, the left hand was pressed down on the cloth at the same time as the right thumb, and the left index finger was also subjected to force. Conversely, there were two peaks of the right thumb in the interval of the second operation (approximately 20,000 ms) and the third operation (approximately 33,000 ms). This is a scene in which the lifting and dropping operations of the presser foot were performed consecutively. In this case, the first strain change of the right thumb and the second strain change of the left index finger occurred simultaneously. This means that the lifting operation and the cloth rotation operation started at almost the same time (Fig. 8 (b)).

In the sewing section, the strain outputs of the right index finger, right middle finger, and left index finger were constantly changing. This is because the index and middle fingers of both

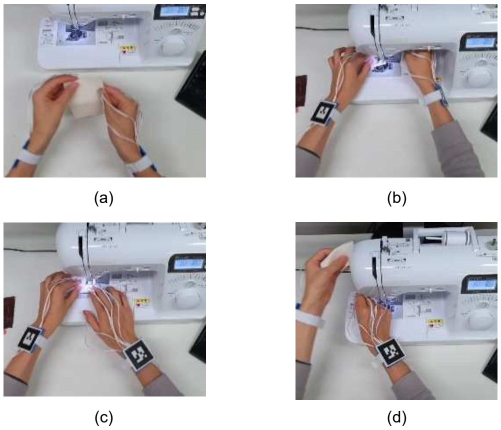



Fig. 8 Hand posture of sewing work

hands were primarily used to press the cloth onto the sewing table, and each finger repeatedly contacted the cloth. Conversely, the output of the right thumb barely changed. The image shown in Fig. 8 (c) also shows that the right thumb did not manipulate the cloth or the sewing machine.

In the cut yarn tails section, as depicted in Fig. 8 (d), the left hand held the fabric, and the right hand held the threads; then, the cutter on the left side of the sewing machine was used to cut the threads. Two peaks of the left index finger appear at 50,000 ms in Fig. 7, but this is because the thread was not cut in the first cutting operation and the cutting operation was redone.

## VII. Conclusions

In this study, we proposed a system for measuring hand posture and fingertip force for training unskilled workers who sew with sewing machines. The system measures hand posture and fingertip force without interfering with the operator's fingertip sensation. Furthermore, we classified sewing behaviors on the basis of color image sequences, hand posture, and fingertip force. Finally, the measured fingertip force data were used to visualize the sewing operator's skills. The classification accuracy of the method using time-series images as input is high; thus, it is suitable to use in work process recognition. Conversely, detailed timing of applying force and interaction with other fingers were confirmed for the fingertip force. Therefore, it is expected to extract skilled techniques by comparing the work measurement results of skilled and unskilled workers.

In the future, we will construct a teaching system for unskilled sewing workers based on the measurement and classification system proposed in this study. Furthermore, we will extract sewing skills by analyzing the visualized sewing skills.

## Acknowledgment

This work was partially supported by KAKENHI and NEDO.

## References

[1] T. Kitamura, H. Teshima, D. Thomas, H. Kawasaki, "Refining OpenPose with a new sports dataset for robust 2D pose estimation," in IEEE/CVF Winter Conference on Applications of Computer Vision Workshops (WACVW), pp.672-681, 2022.

[2] B. Azadi, M. Haslgrübler, G. Sopidis, M. Murauer, B. Anzengruber , A. Ferscha, "Feasibility analysis of unsupervised industrial activity recognition based on a frequent micro action," in 2019 ACM International Conference on Pervasive Technologies Related to Assistive Environments, pp368-375, 2019.

[3] J. Malc and U. Martincz-Hcrnandcz, "Recognition of human activity and the state of an assembly task using vision and inertial sensor fusion methods," in IEEE International Conference on Industrial Technology (ICIT), pp.919-924, 2021.

[4] C. Chengjun, W. Tiannuo, L. Dongnian, H. Jun, "Repetitive assembly action recognition based on object detection and pose estimation," in Journal of Manufacturing System, vol. 55, pp.325-333, 2020.

[5] B. Azadi, M. Haslgrübler, G. Sopidis, M. Murauer, B. Anzengruber, A. Ferscha, "Feasibility analysis unsupervised industrial activity recognition based on a frequent micro action," in Proceedings of the 12th ACM International Conference on Pervasive Technologies Related to Assistive Environments, pp.368-375, 2019.

[6] A. Riedel, N. Brehm, T. Pfeifroth, "Hand Gesture Recognition of Methods-Time Measurement-1 Motions in Manual Assembly Tasks Using Graph Convolutional Networks," in Applied Artificial Intelligence, vol. 36, 2022.

[7] Y. Mo, Y. Nakagawa, K. Nagahama and K. Yamazaki, "Motion Measurement and Segmentation Toward Automated Sewing Operations", 2021 IEEE International Conference on Mechatronics and Automation (ICMA), pp.790-796, 2021.

[8] C. Fermüller, F. Wang, Y. Yang, K. Zampogiannis, Y. Zhang, F. Barranco, M. Pfeiffer, "Prediction of Manipulation Actions," International Journal of Computer Vision, 126(2), pp.358-374, 2018.

[9] V. Becker, P. Oldrati, L. Barrios, G. Sörös, "Touchsense: classifying finger touches and measuring their force with an electromyography armband," in Proceedings of the 2018 ACM International Symposium on Wearable Computers, pp.1-8, 2018.

[10] PyTorch, https://pytorch.org/, (viewd on Friday 14, 2023)

[11] D. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," International Conference on Learning Representation (ICLR), pp. 1-15, 2015.

[12] MediaPipe Hands, https://google.github.io/mediapipe/solutions/hands, (viewed on February 23, 2023)

[13] K. Yamazaki, Y. Nakagawa, A. Ishikawa and M. Hirayama, "Force Sensing Based on Nail Deformation for Measurement of Fingertip Force in Detailed Work," International Conference on Automation Science and Engineering (CASE), pp.2415-2420, 2022.

[14] P. Vinayavekhin, S. Kudoh, J. Takamatsu, Y. Sato, K. Ikeuchi, "Representation and Mapping of Dexterous Manipulation through Task Primitives," in 2013 IEEE International Conference on Robotics and Automation (ICRA), pp.3722-3729, 2013.

[15] E. Ho and T. Komura, "Indexing and Retrieving Motions of Characters in Close Contact," in IEEE Transactions on Visualization and Computer Graphics, vol. 15 No.3, pp.481-492, 2009.