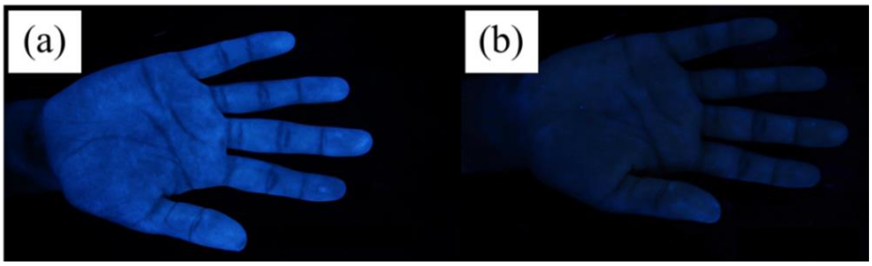

Fig. 5. Images taken under UV light (a) before handwashing and (b) after handwashing.

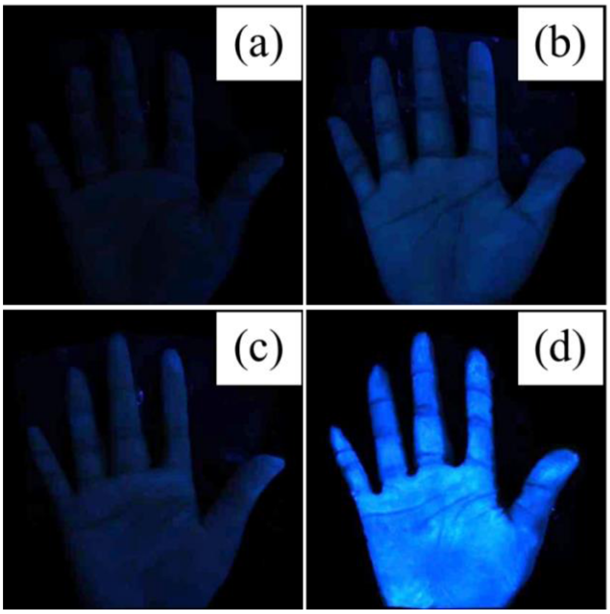

Fig. 6. Examples of classification results in segment 5. (a) The label and CNN prediction are good (b) the label is good and the CNN prediction is bad (c) the label is bad and the CNN prediction is good (d) the label and the CNN prediction are bad.

shows photos taken before handwashing in visible light and UV light. The photos used for deep learning were preprocessed in three steps before being inputted. First, the lower part of the wrists was excluded from the obtained photos and only the palms were extracted. After this, image registration using the Demons' algorithm was applied to all palm shapes photos taken after hand washing [8]. A typical example of image registration using the Demons' algorithm is shown in Fig. 3. Finally, an associate professor from the School of Nursing created a correct response label for all hand washing images.

3. Convolutional Neural Networks (CNN)

Recently, deep learning has been applied in various fields [9,10]. One method of deep learning is the convolutional neural network (CNN), a network often used for image analysis that has good performance results [11,12]. CNNs allow not only a correct answer but also a multiple labeling method with many correct answers [13,14]. Our proposed CNN was also constructed based on multiple labeling methods. As the proposed CNN used supervised learning, it was necessary to attach labels to the data of the hand images taken under UV light. The image data used in this study were classified into two classes (good and bad) for each area by an associate professor from the School of Nursing. Figure 4 shows how the hands were divided into each segmentation. The 585 images collected were randomly classified into 535 training data and 50 test data. Of the 585 images, the number of labels judged to be as good in each part was 86, 74, 78, 66, 117, and 79. The CNN structure with the classification was composed of four convolutions, two max-pool, and three full-connected layers. In addition, the Softmax classification function was applied. The CNN hyperparameters were determined by Bayesian optimization [15] within the range shown in Table 1. Finally, the proposed CNN performance was evaluated by 9-fold cross-validation.

4. Results

Figure 5 shows images taken under UV light before and after handwashing. The pre-wash image shows how the fluorescent paint applied to the hands glows under the blacklight (Fig. 5a). The image after handwashing reveals an absence of fluorescence as the fluorescent paint has been effectively removed after handwashing (Fig. 5b). Table 2 shows the CNN architecture configured based on the hyperparameters determined by Bayesian optimization. The CNN architecture using the determined hyperparameter was verified with 9-fold cross-validation. As a result, the F-value was found to be $81.26 \pm 6.18\%$. Table 3 shows the F-value of each segmentation. Notably, the best F-value was $85.47 \pm 7.12\%$ in segmentation 4 and the worst F-value was $77.44 \pm 6.79\%$ in segmentation 1. Figure 6 shows examples of classification results at segmentation 5. Figure 6a shows that both the label and the CNN prediction are good in all segmentations; Figure 6b shows that the label is good and the CNN prediction is bad; Figure 6c shows that the label in segmentation 1 is bad and the CNN prediction is good; and Figure 6d shows that both the label and the CNN prediction are bad in all segmentations.

5. Discussions

In this paper, we proposed a novel method for quantitative evaluation of handwashing skills based on a convolutional neural network (CNN). Moreover, we created an image capture box equipped with blacklights and took photos of the fluorescent paint applied to the hands before and after handwashing. The resulting F-value was $81.26 \pm 6.18\%$ in the handwashing image evaluation system when CNN was used. Alternatively, considering the F-values of each seg-