Deep Learning-Based Diagnosis of Fatal Hypothermia Using Post-Mortem Computed Tomography

Yuwen Zeng,[1] Xiaoyong Zhang,[2,3] Issei Yoshizumi,[4] Zhang Zhang,[1]
Taihei Mizuno,[5] Shota Sakamoto,[5] Yusuke Kawasumi,[4] Akihito Usui,[4] Kei Ichiji,[4]
Ivo Bukovsky,[6,7] Masato Funayama[4] and Noriyasu Homma[4]

[1]Department of Intelligent Biomedical Systems Engineering, Graduate School of Biomedical Engineering, Tohoku University, Sendai, Miyagi, Japan
[2]National Institute of Technology, Sendai College, Sendai, Miyagi, Japan
[3]Institute of Development, Aging and Cancer, Tohoku University, Sendai, Miyagi, Japan
[4]Department of Radiological Imaging and Informatics, Tohoku University Graduate School of Medicine, Sendai, Miyagi, Japan
[5]Department of Management Science and Technology, Graduate School of Engineering, Tohoku University, Sendai, Miyagi, Japan
[6]Faculty of Science, University of South Bohemia in Ceske Budejovice, Ceske Budejovice, Czech Republic
[7]Faculty of Mechanical Engineering, Czech Technical University in Prague, Prague, Czech Republic

In forensic medicine, fatal hypothermia diagnosis is not always easy because findings are not specific, especially if traumatized. Post-mortem computed tomography (PMCT) is a useful adjunct to the cause-of-death diagnosis and some qualitative image character analysis, such as diffuse hyperaeration with decreased vascularity or pulmonary emphysema, have also been utilized for fatal hypothermia. However, it is challenging for inexperienced forensic pathologists to recognize the subtle differences of fatal hypothermia in PMCT images. In this study, we developed a deep learning-based diagnosis system for fatal hypothermia and explored the possibility of being an alternative diagnostic for forensic pathologists. An in-house dataset of forensic autopsy proven samples was used for the development and performance evaluation of the deep learning system. We used the area under the receiver operating characteristic curve (AUC) of the system for evaluation, and a human-expert comparable AUC value of 0.905, sensitivity of 0.948, and specificity of 0.741 were achieved. The experimental results clearly demonstrated the usefulness and feasibility of the deep learning system for fatal hypothermia diagnosis.

Introduction

Fatal hypothermia occurs when the body's core temperature drops below the normal range due to exposure to extreme cold and the consequential cardiovascular and respiratory failure. It is one of the main causes of mortality in cold climates and accidents such as mountaineering. The diagnosis of hypothermia relies on a combination of autopsy findings related to hypothermia and the exclusion of other possible causes of death. However, it could be difficult to perform autopsies due to the time cost or culture reasons, especially in countries with low autopsy rates such as Japan. Therefore, post-mortem computed tomography (PMCT) was introduced to assist in the diagnosis of hypothermia, which could provide a non-invasive and comprehensive evaluation of the body's internal structures and abnormality information (Michiue et al. 2012).

With the assistance of PMCT, Kawasumi et al. (2013) found that many cases of hypothermic death exhibited certain characteristics, including a lack of increased concentration in the lung-field, blood clotting in the heart, thoracic aorta, or pulmonary artery. Hyodoh et al. (2016) investi-

gated the time-related course of post-mortem lung changes on rabbits and found the percent of aerated lung volume remained significantly high in hypothermia. Other studies (Hyodoh et al. 2013; Schweitzer et al. 2014) reported that hypothermia was associated with significantly lower lung PMCT attenuation and lung weights. However, these findings have a relatively low specificity because researchers found some causes, like carbonic oxide poisoning and severe diabetic ketoacidosis, also correlating with below-normal PMCT lung attenuation (Kawasumi et al. 2013; Schweitzer et al. 2014). But most of all, there is a shortage of forensic pathologists (Weedn and Menendez 2020), and even fewer can interpret radiology images. The accuracy of the diagnosis may also vary depending on the forensic pathologist's proficiency.

To overcome the challenges above, we proposed the first deep learning-based diagnosis system for hypothermia, and further explored the possibility of utilizing the models for direct confirmation after the on-site investigation, instead of just helping with diagnosis. Deep learning (DL) has shown remarkable performance in classifying and detecting various pathologies in medical images. It can automatically learn representative features from raw data without requiring handcrafted feature in traditional methods. For example, DL-based diagnosis systems were developed for drowning and showed high performance on classification (Homma et al. 2020; Sakamoto et al. 2021; Zeng et al. 2021, 2023). They showed the feasibility of delivering accurate diagnosis using DL-based systems along with information from the on-site investigation.

In this paper, we trained deep convolutional neural networks (DCNNs) to diagnose fatal hypothermia using PMCT images and evaluated them on an independent test set. In order to validate the effectiveness of the DCNNs and to provide transparent results, we visualized areas that were important to the models and discussed some typical cases.

Materials and Methods

Data and imaging conditions

As part of pre-autopsy screening, PMCT scanning was performed on a multi-channel scanner (Canon Medical Systems, Otawara, Japan). The same type of scanner was used throughout the study, ensuring consistency in the imaging parameters and quality. Chest CT images of size 512×512 pixels were acquired using a 1 mm $\times$ 4-row slice configuration mode. Most cases contained 28 images at seven different levels, from the pulmonary apex to the lowest part of the left lung. Each level was composed of 4 slices, and the number of regions in a case varied from six (24 images) to nine (36 images) due to individual differences, such as stature. There were four forensic pathologists who performed the operation, and the diagnosis of the cause of death was based on a mutual consensus. The final check of autopsy reports was made by a 35-year experienced forensic pathologist to ensure the consistency for all cases (Funayama 2008). In this study, as we aimed to use lung features to diagnose hypothermia with DCNNs, only the former 24 images of each case were used because the images of the lower part mostly present liver and stomach. An example of a PMCT case is shown in Fig. 1.

We selected 108 autopsy cases (64 males and 44 females) diagnosed as hypothermia from the Autopsy Imaging Center, Tohoku University Graduate School of Medicine, from January to December in 2021. One male and two female cases were excluded because the direct causes of death were not hypothermia, despite the low core temperature when dying. The average ages of the remaining males and females were 70.9 (range 33-95) and 74.0 (range 25-93) years, respectively. Meanwhile, we had 115 death cases as the non-hypothermia group (71 males and 44 females) from 2015 to 2021. The average ages of the males and females were 54.5 (range 21-92) and 58.5 (range 21-94) years, respectively. The causes of death in the non-hypothermia group were cardiovascular disease (n = 46), asphyxia (n = 13), infection (n = 7), poisoning (n = 15), trauma (n = 16), alcoholic and diabetic ketoacidosis (n = 6), and other causes like subarachnoid hemorrhage (n = 12). The postmortem time was ranging 0.5-21 days (3.4 ± 3.7, mean ± SD) in hypothermia cases and 0.4-5 days (1.4 ± 0.8) in non-hypothermia cases. Besides, the time of 16 hypothermia and 7 non-hypothermia cases cannot be determined, probably ranging from several days to several weeks. Although the average postmortem time of the hypothermia
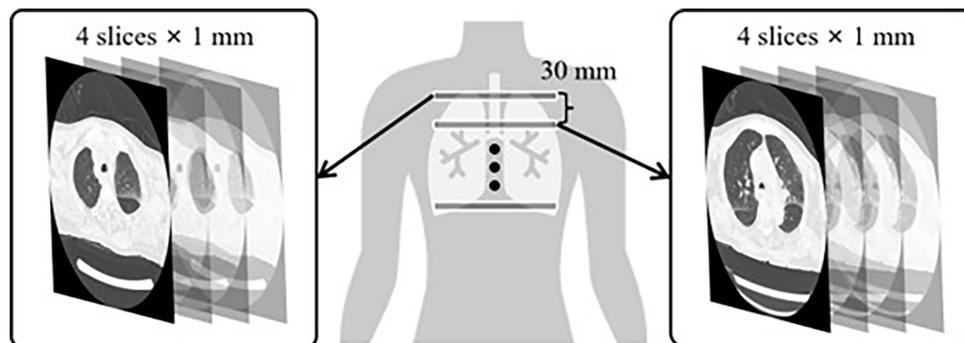

Fig. 1. A simplified example of a post-mortem multi-slice CT case.

cases is longer, due to the scene of low temperature and the outdoors, decomposition in bodies was slow. Also, a significant difference in the age distribution can be observed between two groups as the elderly are more susceptible to hypothermia and are at higher risk of dying from it compared to younger individuals. We empirically excluded cases with advanced decomposition, infants, severe carbonization, drowning, and severe chest trauma in both groups. Among all cases, cardiopulmonary resuscitation (CPR) was performed in eight hypothermia cases and 23 non-hypothermia cases. As mentioned before, fatal hypothermia was associated with significantly lower lung PMCT attenuation and lower lung weights, so we also listed the lung weights (g) in both hypothermia and non- hypothermia cases (control) in Table 1.

The use of PMCT images for this study was approved by the ethics board of Tohoku University (protocol number: 2021-1-495; date: 2018-09-18). Informed consent was not required for this research.

Deep convolutional neural network

Considering different DL model has its own strengths, and the most significant characteristic of fatal hypothermia is low lung PMCT attenuation (not feature-rich for computer vision), we chose to compare three mainstream DCNN architectures that have achieved state-of-the-art results on image classification. The three models were Inception-V3 (Szegedy et al. 2016), VGG-16 (Simonyan and Zisserman 2014), and ResNet18 (He et al. 2016).

Inception-V3 uses a combination of parallel convolutions with different filter sizes to capture features at different scales. It has 48 layers and includes multiple Inception modules, which are composed of parallel convolutions with different filter sizes and pooling layers. VGG-16 contains 16 layers, including 13 convolutional layers and three fully connected layers. The architecture is known for its simplicity and uniformity. ResNet18 has skip connections to address the problem of vanishing gradients in very deep neural networks. It has 18 layers, including 16 convolutional layers and two fully connected layers. VGG-16 is known for its simplicity and uniformity, with all convolutional layers having a small 3×3 filter size and the same padding and stride. ResNet18 and Inception-V3 use their unique architectures to address the problem of vanishing gradients in deep neural networks.

To make the models converge faster, we tuned on the ImageNet-pretrained models (Deng et al. 2009), and substituted two new fully connected layers for the original layers on top of the last convolutional layers. The outputs of models provided a probability for each group: hypothermia or non-hypothermia (control group). A higher probability indicated a greater likelihood of the input being hypothermia, and vice versa. The final determination of a hypothermia or non-hypothermia was based on whether the probability was greater or less than a default threshold of 0.5. When the probability is close to 0.5, it suggested the input might exhibit both characteristics of each class and could be a hard example to classify.

Table 1. Lung weights (g) in the hypothermia and non-hypothermia cases.

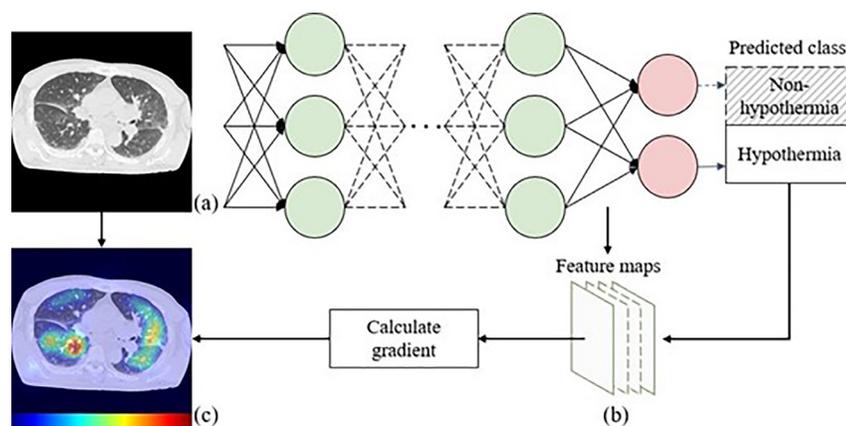| Lung	Hypothermia		Non-hypothermia	
	Male	Female	Male	Female
Left	314.2 ± 91.2	236.6 ± 91.9	548.8 ± 158.9	390.8 ± 184.9
Right	365.2 ± 107.0	282.7 ± 108.1	649.8 ± 195.1	460.9 ± 222.4

Data are shown as mean ± SD.

Fig. 2. An illustration of the visualization method.
Given an input image (a), we compute the gradient of the score for a target class with respect to feature maps of the last convolutional layer (b). This gradient is then further processed and used to weight the feature maps and overlap with the input to obtain the visualization result (c).

We randomly split the original PMCT dataset into the training and test set, each containing 93 and 15 hypothermia cases (2,188 and 404 images), and 100 and 15 non-hypothermia cases (2,416 and 428 images), respectively. In a corresponding manner, the results were given in image-wise and case-wise, where the case-wise results were obtained by averaging all predictions of a same case. To reduce overfitting on the small dataset, we applied data augmentation on the training set, which will not increase the total sample size (Shorten and Khoshgoftaar 2019). During the training process, an input would be randomly transformed into one of the following states: unchanged, horizontally/vertically flipped, flipped, horizontally/vertically shifted in the range of 20% of the width/height of the input, or randomly rotated in the range of 20°. All models were optimized using the Adam optimizer with a learning rate of 1×10^{-6} and a batch size of eight.

Visualization

DCNNs can learn to recognize complex patterns in images, but their internal workings are often opaque. This makes it difficult to understand which region of the input image was important to the prediction (Singh et al. 2020; Tjoa and Guan 2021). In order to evaluate the effectiveness of the DCNNs and to provide transparent results, we utilized a saliency visualization technique called Gradient-weighted Class Activation Mapping (Grad-CAM) (Selvaraju et al. 2017), which is used in computer vision to understand which parts of an image a DCNN focuses on while making a prediction.
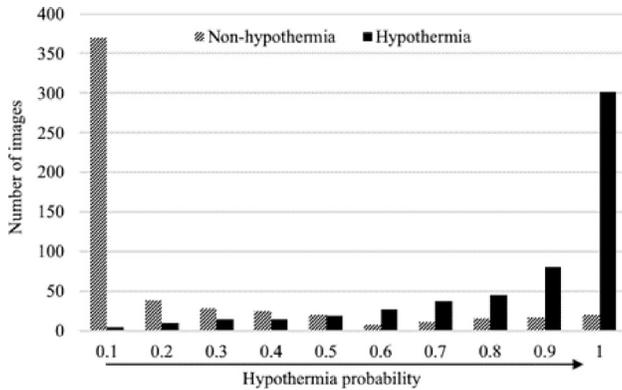

Fig. 3. The probability distribution of the image-wise predictions for hypothermia and non- hypothermia group in test set.

A simple illustration of Grad-CAM is shown in Fig. 2. To generate the heatmap from a given input (a), we can calculate the gradient (b) of the final convolutional layer of the network with respect to the predicted class. The gradients are then global average-pooled to obtain the importance of each output of the final convolutional layer, i.e., feature map. Finally, the weighted combination of feature maps is used to overlap with the input to generate the visualization
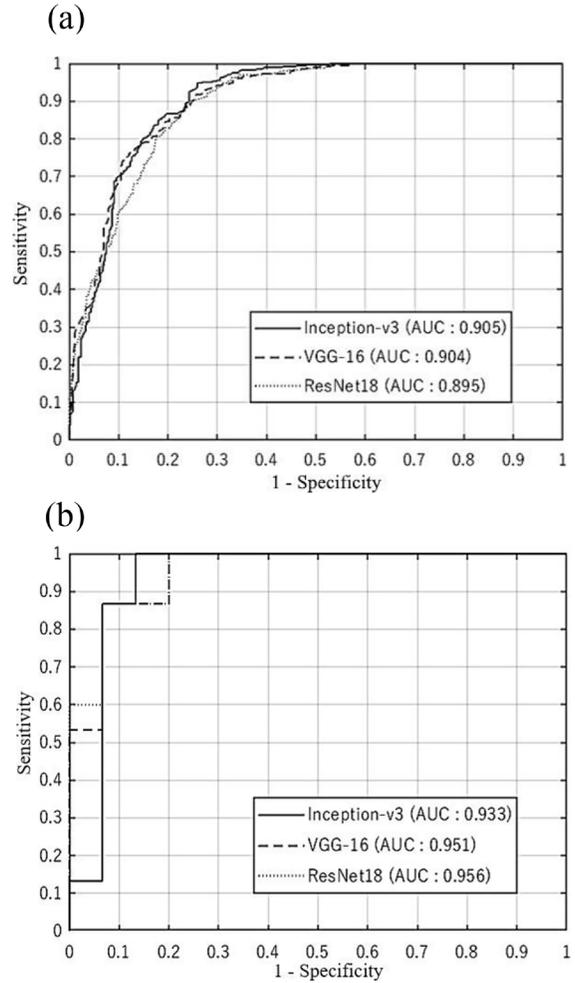
(a)

(b)

Fig. 4. The receiver operating characteristic (ROC) curves and the area under the ROC curve (AUC) values of three models (Inception-V3, VGG-16 and ResNet18). (a) Image-wise results. (b) Case-wise results. The numeric variables in ROCs are the hypothermia probability predicted by models.

Table 2. The sensitivity, specificity and AUC of models on the test set.

Models	Image-wise			Case-wise		
	Sensitivity	Specificity	AUC	Sensitivity	Specificity	AUC
Inception-V3	0.948	0.741	0.905	1	0.867	0.933
VGG-16	0.918	0.743	0.904	0.867	0.933	0.951
ResNet18	0.884	0.769	0.867	0.867	0.933	0.956

AUC, the area under the ROC curve.

(a)

Prob: 0.866   Prob: 0.802   Prob: 0.983   Prob: 0.788

(b)

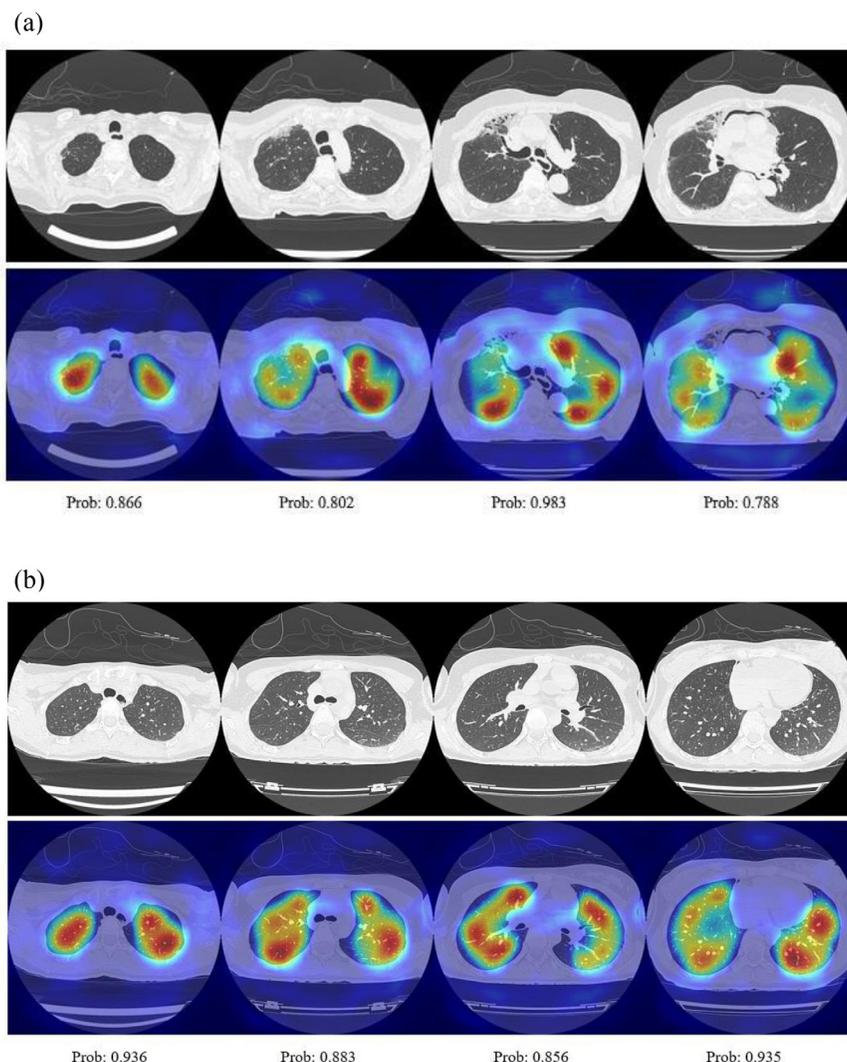Prob: 0.936   Prob: 0.883   Prob: 0.856   Prob: 0.935

Fig. 5. Two true positive (TP) cases.
The upper row is the original images, and the bottom is the visualization with the probabilities of hypothermia. (a) Case 1, female, aged 77. (b) Case 2, female, aged 25.

result (c). This whole process would provide a visual explanation for the network's decision, indicating the regions of the input image that influenced the output. Warmer (red) regions correspond to higher scores for the target class, meaning these areas are more important to the model.

Results

A default threshold of 0.5 was used to predict the input into hypothermia or non-hypothermia, then the distribution of predicted hypothermia probabilities can be presented as Fig. 3. Although a small proportion of images would be misclassified as false positive (FP) or false negative (FN) cases, most images can be correctly predicted with a high confidence as depicted in both ends of the horizontal axis.

We used the receiver operating characteristic (ROC) curve, and area under the ROC curve (AUC) to evaluate the models' performance on hypothermia prediction, as shown in Fig. 4. Based on the image-wise predictions (a), we calculated the case-wise predictions (b) as well. All images contained in most cases were correctly classified, but in a small number of cases there were several misclassified images. By averaging the probabilities of all images, we can further improve the classification performance of models. The sensitivity, specificity, and AUC of the three models are given in Table 2. Inception-V3 achieved the best image-wise performance with a sensitivity of 0.948 and AUC of 0.905, and a case-wise sensitivity of 1 and AUC of 0.933. The image-wise specificity was lower than their sensitivity for all models, but the situation conversed on the case-wise results of VGG-16 and ResNet18. Such change might be caused by the averaging operation on image-wise results, as it could change the optimal operating point on the ROC curves.

(a)

Prob: 0.167   Prob: 0.084   Prob: 0.005   Prob: 0.016

(b)

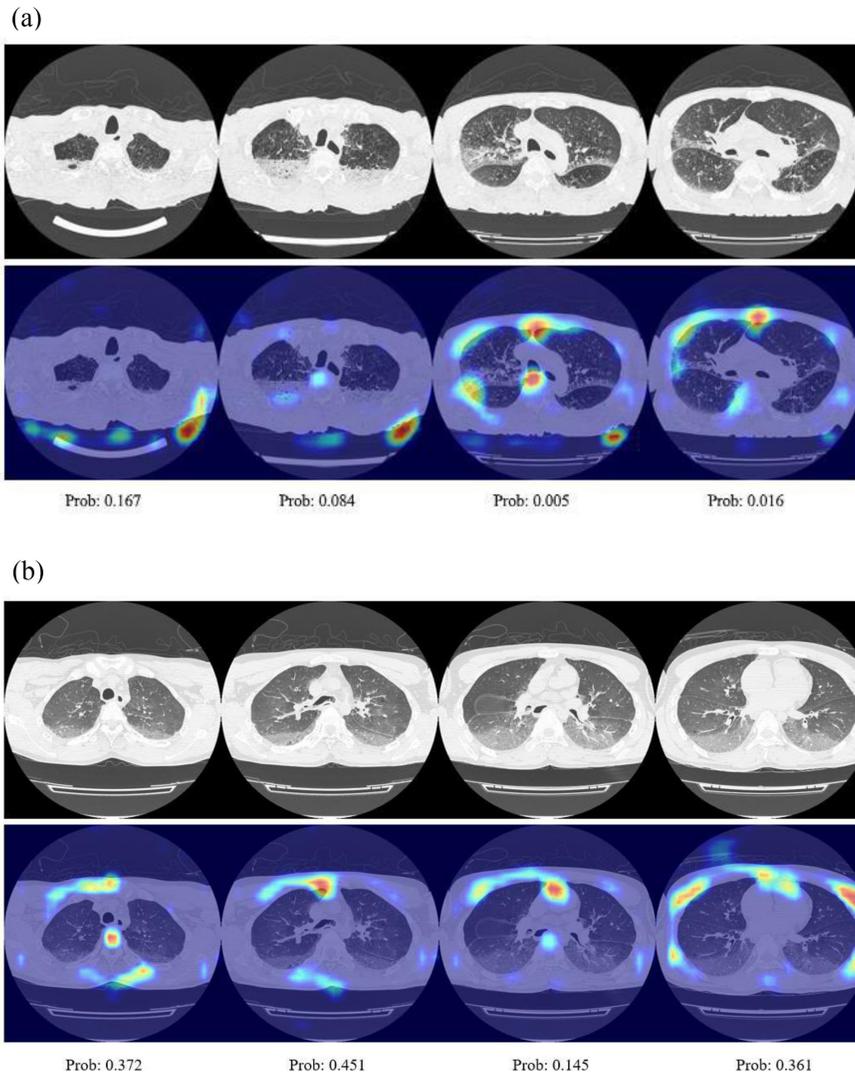Prob: 0.372   Prob: 0.451   Prob: 0.145   Prob: 0.361

Fig. 6. Two true negative (TN) cases.
The upper row is the original images, and the bottom is the visualization with the probabilities of hypothermia. (a) Case
3, male, aged 64. (b) Case 4, male, aged 35.

Discussion

To understand how the models make decisions and provide insight into their inner workings, we chose VGG-16 as an example and visualized the important regions of an input image that models use to make their prediction. Considering the difference in the age distribution of hypothermia and non-hypothermia cases, we gave one case of an elder person and one case of a younger person for both true positive (TP) and true negative (TN) examples. All images in each case were classified correctly.

The TP examples were shown in Fig. 5. Case 1 (female, aged 77; Fig. 5a) had a left and right lung weight of 210 g and 360 g. The shadow on the ventral mediastinum side might be due to the lower right prone position when the body was found. It can be observed that the model focused on the low lung attenuation area, which is consistent with the findings we mentioned before. Case 2

(female, aged 25; Fig. 5b) had a left and right lung weight of 250 g and 200 g, also showing low lung attenuation. Similar to Case 1, the model paid attention to nearly the whole lung area.

The TN examples were shown in Fig. 6. The cause of death of Case 3 (male, aged 64; Fig. 6a) was bronchopneumonia and underwent cardiopulmonary resuscitation (CPR), with left lung weight of 630 g and right lung weight of 730 g. Case 4 (male, aged 35; Fig. 6b) was diagnosed as acute circulatory failure and received CPR from his family, with left lung weight of 510 g and right lung weight of 590 g. Similar to Case 1, the model paid attention to nearly the whole lung area. The model barely considered the lung opacity as a characteristic of hypothermia, and correctly predicted most images in both cases with low hypothermia probabilities. Some predictions were of a probability close to 0.5, suggesting the low confidence of the model.

In the following, we discussed some misclassification

(a)

Prob: 0.676   Prob: 0.682   Prob: 0.732   Prob: 0.970

(b)

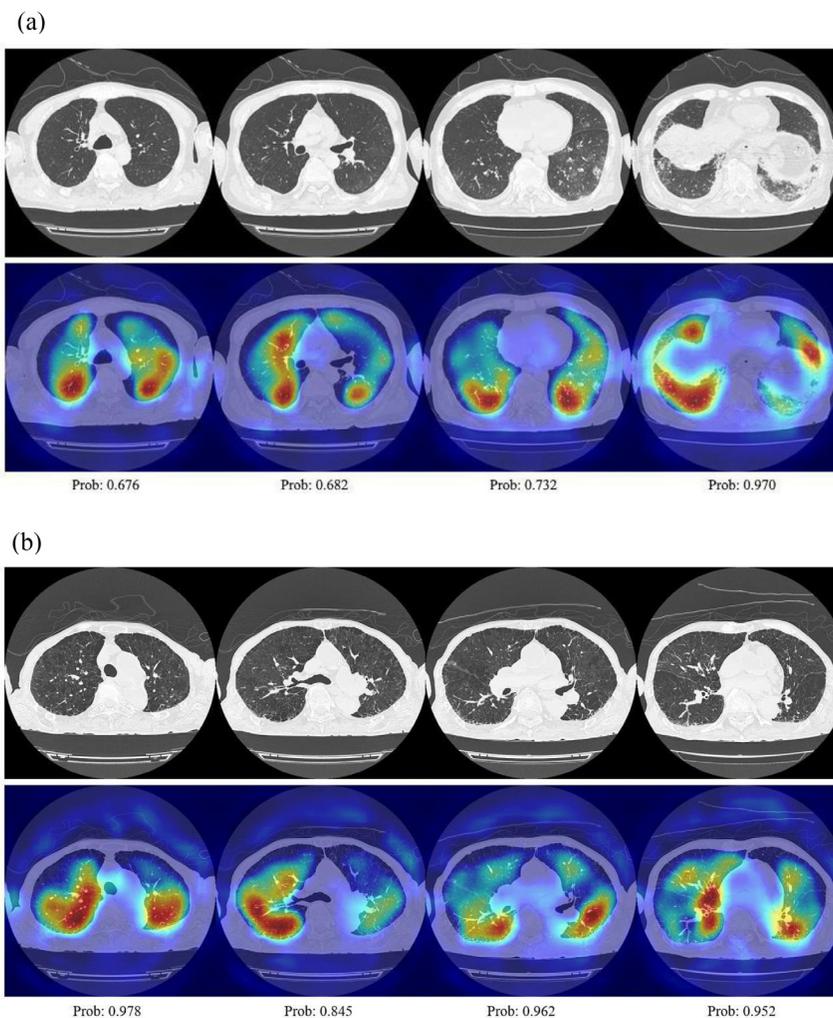Prob: 0.978   Prob: 0.845   Prob: 0.962   Prob: 0.952

Fig. 7. Two false positive (FP) cases.
The upper row is the original images, and the bottom is the visualization with the probabilities of hypothermia.
All images in each case were misclassified. (a) Case 5, male, aged 77. (b) Case 6, male, aged 91.

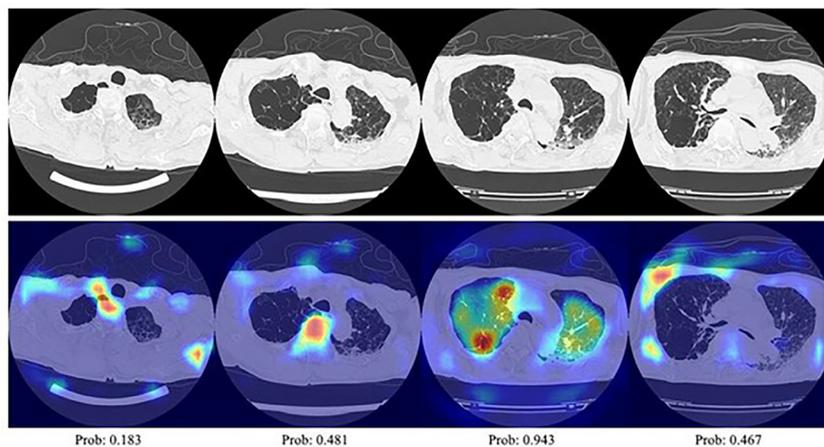Prob: 0.183   Prob: 0.481   Prob: 0.943   Prob: 0.467

Fig. 8. A false negative (FN) case.
The upper row is the original images, and the bottom is the visualization with the probabilities of hypothermia. There
were 15 out of 24 images in this case being misclassified as non-hypothermia. Case 7, male, aged 84.

in the test set. Among the 15 non-hypothermia cases, two cases (including all images) were completely misclassified as hypothermia cases (FP). Case 5 (male, aged 77; Fig. 7a) was bronchopneumonia, with left and right lung weight of 370 g and 310 g. Compared with the TN Case 3, we may find they had the same diagnosis of bronchopneumonia, except Case 5 not rederiving CPR. Although there was inflammation in the base of the lung, the other observations were significantly different. Its low lung CT attenuation and lung weights (Table 1) were very similar to the characteristics of hypothermia, which might be the reason being misclassified. Case 6 (male, aged 91; Fig. 7b) was diagnosed as ischemic heart disease by forensic pathologists, with left and right lung weight of 260 g and 330 g. This case presented almost the same features as Case 5 and was also misclassified as a hypothermia case.

Among the 15 hypothermia cases, only one case (Case 7) was misclassified as non-hypothermia cases (FN). As shown in Fig. 8, Case 7 (male, aged 84) had emphysema and undergone CPR. The left and right lung weight was 350 g and 390 g, which was closer to the average of hypothermia cases in Table 1. There were 15 out of 24 images in Case 7 being misclassified as non-hypothermia, in which 12 of them were predicted with a low hypothermia probability near to 0.210 ($\pm$ 0.045). Only four images were predicted with a hypothermia probability higher than 0.8. The above-mentioned made Case 7 a hard example to classify, and even forensic pathologists need to consider other information to make the final decision. Although Case 7 was the only misclassified case in the test set that received CPR (seven in total), it is hard to tell whether the model can accurately diagnose hypothermia-CPR cases since the sample size is small. Because CPR does affect AI diagnosis of fatal hypothermia, near-complete differentiation is difficult at this time and remains a future challenge.

There is some limitation in this study. For Case 6 and 7, the area of the lung parenchyma where the presence or absence of edema can be observed decreases due to emphysema (air cyst). One may notice that the average age of the hypothermia cases was significantly higher than that of the non-hypothermia subjects, and the elderly usually have a higher incidence of emphysema than the young. But in our dataset, there were only two cases with emphysema in hypothermia and one in non-hypothermia. Thus, there is little chance that the results were influenced by the higher incidence of emphysema among the elderly. This potential bias would be seriously considered in future work. Also, these misclassifications show the non-specificity of hypothermia characteristic. The low lung CT attenuation and lung weights, which are characteristic of hypothermia, can also be present in non-hypothermic dehydration cases without pulmonary edema, such as starvation, heat stroke, and blood loss, which cannot be distinguished with hypothermia via postmortem imaging or examination. Meanwhile, it is difficult to collect dehydration cases (other than hypothermia) for the control group. However, combining with the on-site investigation information, we can further ensure the classification results of the model based on the high image-wise sensitivity of 0.948 and AUC of 0.905, and case-wise sensitivity of 1 and AUC of 0.933.

In conclusion, we proposed a deep learning-based computer-aided diagnosis system for hypothermia using post-mortem lung CT images. Three models were trained, compared, and evaluated on an independent test set. Then we discussed the visualization results that provide us with a better comprehension on the decision of models. Through the detailed information from autopsy, we gave in-depth analysis of those misclassified cases and highlighted the usefulness and feasibility of the high-performance DL models. Given there would be a lot of information from on-site investigation about the discovered bodies, we are considering the models' potential to be used for direct confirmation rather than assistance for forensic pathologists. As future work, we can ensemble multiple classifiers to reduce the number of false positives produced by any individual feature. Also, by adjusting the threshold of classification probabilities, the trade-off between sensitivity and specificity can be optimized based on the specific application.

Acknowledgments

Conflict of Interest

The authors declare no conflict of interest.

References

Deng, J., Dong, W., Socher, R., Li, L.J., Li, K. & Fei-Fei, L. (2009) Imagenet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*), Miami, FL, pp. 248-255.

Funayama, M. (2008) Diagnosis of hypothermia. *Nihon Hoigaku Zasshi*, **62**, 145-154 (in Japanese).

He, K., Zhang, X., Ren, S. & Sun, J. (2016) Deep Residual Learning for Image Recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*), Las Vegas, NV, pp. 770-778.

Homma, N., Zhang, X., Qureshi, A., Konno, T., Kawasumi, Y., Usui, A., Funayama, M., Bukovsky, I., Ichiji, K., Sugita, N. & Yoshizawa, M. (2020) A deep learning aided drowning diagnosis for forensic investigations using post-mortem lung CT images. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, **2020**, 1262-1265.

Hyodoh, H., Shimizu, J., Rokukawa, M., Okazaki, S., Mizuo, K. & Watanabe, S. (2016) Postmortem computed tomography findings in the thorax - experimental evaluation. *Leg. Med.* (*Tokyo*), **19**, 96-100.

Hyodoh, H., Watanabe, S., Katada, R., Hyodoh, K. & Matsumoto, H. (2013) Postmortem computed tomography lung findings in fatal of hypothermia. *Forensic Sci. Int.*, **231**, 190-194.

Kawasumi, Y., Onozuka, N., Kakizaki, A., Usui, A., Hosokai, Y., Sato, M., Saito, H., Ishibashi, T., Hayashizaki, Y. & Funayama,

M. (2013) Hypothermic death: possibility of diagnosis by post-mortem computed tomography. *Eur. J. Radiol.*, **82**, 361-365.

Michiue, T., Sakurai, T., Ishikawa, T., Oritani, S. & Maeda, H. (2012) Quantitative analysis of pulmonary pathophysiology using postmortem computed tomography with regard to the cause of death. *Forensic Sci. Int.*, **220**, 232-238.

Sakamoto, S., Zhang, X., Homma, N., Kawasumi, Y., Usui, A., Ogawara, T., Funayama, M., Ichiji, K., Sugita, N. & Yoshizawa, M. (2021) Explainable deep learning-based system for discrimination of drowning in post-mortem lung CT images. In *18th SICE Symposium on Computational Intelligence*, pp. 49-54 (in Japanese).

Schweitzer, W., Thali, M., Giugni, G. & Winklhofer, S. (2014) Postmortem pulmonary CT in hypothermia. *Forensic Sci. Med. Pathol.*, **10**, 557-569.

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D. & Batra, D. (2017) Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. In *2017 IEEE International Conference on Computer Vision* (*ICCV*), Venice, Italy, pp. 618-626.

Shorten, C. & Khoshgoftaar, T.M. (2019) A survey on image data augmentation for deep learning. *J. Big Data*, **6**, 60.

Simonyan, K. & Zisserman, A. (2014) Very deep convolutional networks for large-scale image recognition. *arXiv*, doi: arXiv:1409.1556.

Singh, A., Sengupta, S. & Lakshminarayanan, V. (2020) Explainable deep learning models in medical image analysis. *J. Imaging*, **6**, 52.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. & Wojna, Z. (2016) Rethinking the Inception Architecture for Computer Vision. In *2016 IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*), Las Vegas, NV, pp. 2818-2826.

Tjoa, E. & Guan, C. (2021) A survey on explainable artificial intelligence (XAI): toward medical XAI. *IEEE Trans. Neural Netw. Learn. Syst.*, **32**, 4793-4813.

Weedn, V.W. & Menendez, M.J. (2020) Reclaiming the autopsy as the practice of medicine: a pathway to remediation of the forensic pathology workforce shortage? *Am. J. Forensic Med. Pathol.*, **41**, 242-248.

Zeng, Y., Zhang, X., Kawasumi, Y., Usui, A., Ichiji, K., Funayama, M. & Homma, N. (2021) Deep learning-based interpretable computer-aided diagnosis of drowning for forensic radiology. In *2021 60th Annual Conference of the Society of Instrument and Control Engineers of Japan* (*SICE*), Tokyo, pp. 820-824.

Zeng, Y., Zhang, X., Kawasumi, Y., Usui, A., Ichiji, K., Funayama, M. & Homma, N. (2023) A 2.5D deep learning-based method for drowning diagnosis using post-mortem computed tomography. *IEEE J. Biomed. Health Inform.*, **27**, 1026-1035.