

Segmentation Method for Face Modelling in Thermal Images

Albar^{a, 1}, Hendrick^{a, 2, *}, Rahmat Hidayat^{b, 3}

^a Department of Electrical Engineering, Politeknik Negeri Padang
Jl. Kampus, Limau Manis, Kec. Pauh, Kota Padang, Sumatera Barat 25162, Indonesia

^b Department of Information Technology, Politeknik Negeri Padang
Jl. Kampus, Limau Manis, Kec. Pauh, Kota Padang, Sumatera Barat 25162, Indonesia

¹ albar@pnp.ac.id; ² hendrick@pnp.ac.id*; ³ rahmat@pnp.ac.id
* corresponding author

ARTICLE INFO

Article history:

Received 01 December 2020

Revised 12 December 2020

Accepted 25 December 2020

Published online 31 December 2020

Keywords:

Face Detection

Segmentation

Thermal Images

Deep Learning

ABSTRACT

Face detection is mostly applied in RGB images. The object detection usually applied the Deep Learning method for model creation. One method face spoofing is by using a thermal camera. The famous object detection methods are Yolo, Fast Region Based Convolutional Neural Networks (RCNN), Faster RCNN, SSD, and Mask RCNN. We proposed a segmentation Mask RCNN method to create a face model from thermal images. This model was able to locate the face area in images. The dataset was established using 1600 images. The images were created from direct capturing and collecting from the online dataset. The Mask RCNN was configured to train with 5 epochs and 131 iterations. The final model predicted and located the face correctly using the test image.

This is an open access article under the CC BY-SA license
(<https://creativecommons.org/licenses/by-sa/4.0/>).

I. Introduction

Face Recognition has been applied in many areas, especially in the security system. Avoiding the spoofing face [1], usually, stereo cameras have been applied to face recognition systems [2]. Nowadays, face recognition is supported by a deep learning method, which is reduced machine learning procedures [3]. When applying machine learning methods, feature extraction is done manually and trained to get the models. By using the deep learning methods, the creation model achieved high accuracy in prediction [4]. The other method to identify the real person is by using a thermal camera which records the subject temperature. Thermal is mostly applied in contactless temperature measurements, such as the steel industry. The thermal camera is not only applied in the industrial field, but also applied in a biomedical application, such as contactless breath rate measurement [5][6], breast health, musculoskeletal, neurological medicine, dermatology, and dental care [7].

Convolutional Neural Network (CNN) is the common method in deep learning which has been applied in many areas such as in biomedical images [8]. Deep Learning has some famous frameworks, which are Tensorflow, Keras, PyTorch, Caffe, CNTK, and MXNet [4]. Region Based Convolutional Neural Networks (RCNN) is one of the best methods for object detection. RCNN has been applied in some applications which are finding optical nerve in fundus images [9], face detection in RGB images [10], and facial detection [11]. In this research, we proposed a segmentation method, Mask RCNN, to create a face model from thermal images. The model will detect and locate the face from thermal images.

Face images were recorded by using a FLIR Lepton thermal camera, which has specification as a military standard device [12]. The dataset was created by combining direct recorded images and images from the online dataset. The dataset was expanded using the data augmentation method to achieve accurate prediction models [13]. The face model was created by using the segmentation method of Mask RCNN. The Mask RCNN is covered by TensorFlow-GPU and Keras framework [14]. To reduce training time, TensorFlow-GPU is applied in this research. The final model was

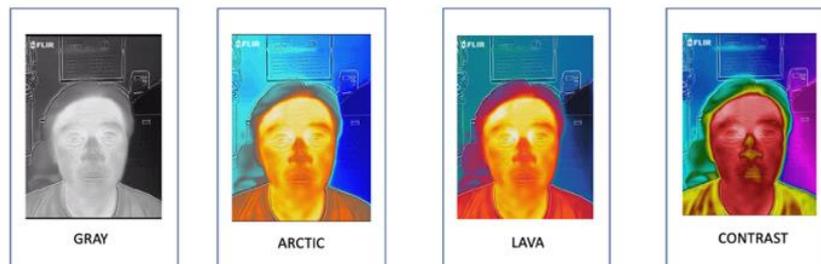


Fig. 1. Thermal image format

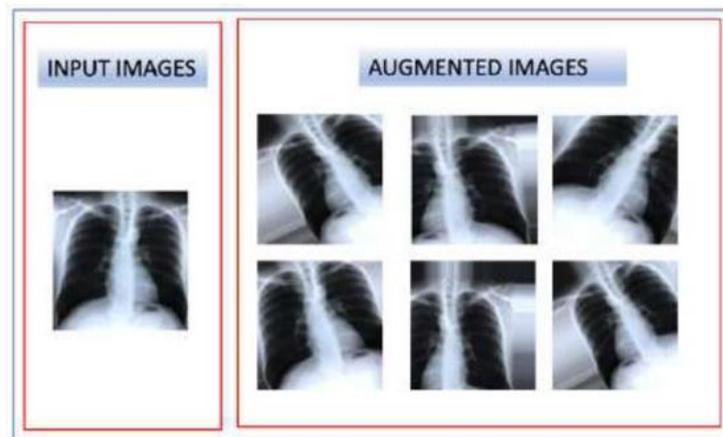


Fig. 2. Augmentation result of the chest X-Ray image

applied in real-time detection by using OpenCV. In future works, this model will be embedded in mini PC, such as Raspberry Pi. This model will be developed to measure face temperature from thermal images.

II. Methods and Materials

A. Data Collection

Data collection was done by using FLIR Lepton Thermal Camera. Data collection is not only from Flir Lepton Thermal images, but also collected from online dataset. The thermal images have some formats which are contrast, gray, arctic, and lava. Each format has their own function. In this research, we only selected contrast format for all dataset. Figure 1 shows the thermal images formats.

To increase the dataset size, image augmentation was applied to the original dataset. Image augmentation could be done by rotating, flipping, etc. The data augmentation methods will create 100 images from each original image. The final dataset size is 1600 images. Figure 2 shows the image augmentation result of the chest X-Ray image [15].

B. Training Preparation

For object detection purposes, every object has to label indicating the object location in images. Label creations were done by using labeling. The final files were saved in XML files. The system was trained and deploy in Ubuntu 16.04 operating system. The dataset was separated into images train and annots folder. Labeling was installed by using a command.

Figure 3 Shows the XML result, which contained object location, image size, image deep. The red square box shows the object location.

C. Mask RCNN

Mask RCNN was a development model of RCNN and fast RCNN. The Fast RCNN produced a class label and a bounding box offset for every candidate object. Mask RCNN has the same output as well as RCNN, but it also created the object Mask. The other important thing that made Mask RCNN

```

<annotation>
  <folder>thermal</folder>
  <filename>1picture-1.jpg</filename>
  <path>/Users/hendrickhendrick/Documents/Research/thermalImages/thermal/1picture-1.jpg</path>
  <source>
    <database>Unknown</database>
  </source>
  <size>
    <width>1080</width>
    <height>1440</height>
    <depth>3</depth>
  </size>
  <segmented>0</segmented>
  <object>
    <name>face</name>
    <pose>Unspecified</pose>
    <truncated>0</truncated>
    <difficult>0</difficult>
    <bndbox>
      <xmin>93</xmin>
      <ymin>339</ymin>
      <xmax>820</xmax>
      <ymax>1112</ymax>
    </bndbox>
  </object>
</annotation>

```

Fig. 3. Object location in XML file

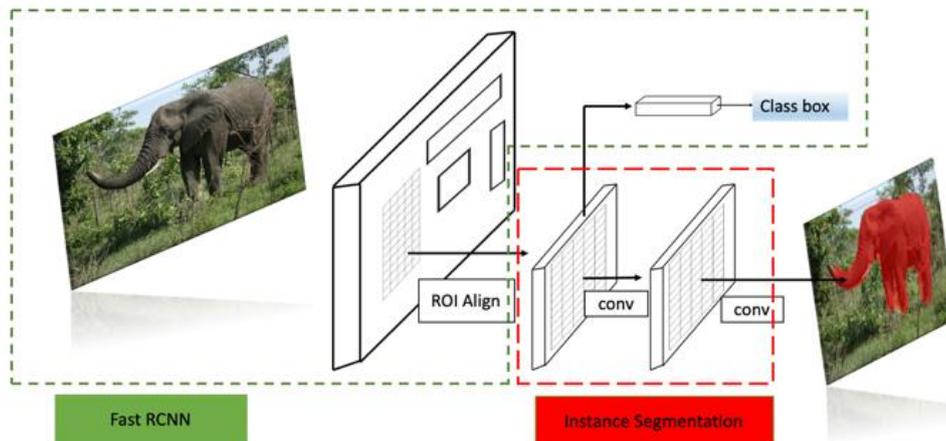


Fig. 4. Mask RCNN framework with instance segmentation

better than RCNN is pixel-to-pixel alignment. ROI Alignment has the function of creating a small feature map for each RoI. The final stage of Mask RCNN is an instance of segmentation. An instance segmentation generates a pixel-wise mask for each object in the image. Even though two objects are in the same class, Mask RCNN treats them as a different instance. Figure 4 shows the framework mask RCNN with instance segmentation.

Training Mask RCNN in python required some libraries to install correctly, especially for CUDA and CUDNN. In this training, we used CUDA 9.0 with Nvidia driver 384. We must consider the laptop specification to decide the version of CUDA, and CUDNN. TensorFlow-GPU and Keras install to the device. It generates an error core dump if the installation is not proper. In this research, we used a laptop which has a specification, as mention in Table 1. The thermal image dataset was divided into train datasets and test datasets. The train dataset size is 80% of all datasets, and the test set is 20%.

Table 1. Laptop specifications

No	Device	Specification
1	CPU	Core i7
2	GPU	GTX 750
3	RAM	16 GByte

III. Results and Discussions

Figure 5 shows the result of image augmentation process. The images became 100 images from each source image. The total images for the dataset are 1600. The training set was configured with epoch = 5 and iteration =131. By setting 5 for the epoch value, the training loop ended at 5 the epoch. After 6 hours, the model was created in h5 formats. Figure 6 shows the created model from training. Because we configured 5 for the epoch value, the model was also created as mask_rcnn_cfg_0001.h5 for 1st epoch, mask_rcnn_cfg_0002.h5 for 2nd epoch, mask_rcnn_cfg_0003.h5 for 3rd epoch, mask_rcnn_cfg_0004.h5 for 4th epoch, and mask_rcnn_cfg_0005.h5 for 5th epoch. All models were saved automatically by training program. To find out the performance, each model was tested by using test dataset. Figure 7 depicts the test image which deployed mask_rcnn_cfg_0005.h5 in the program. The face was predicted perfectly by using the face model. The program automatically created red rectangle to visualize the detected face in thermal images. This stage tested a model with a single image.

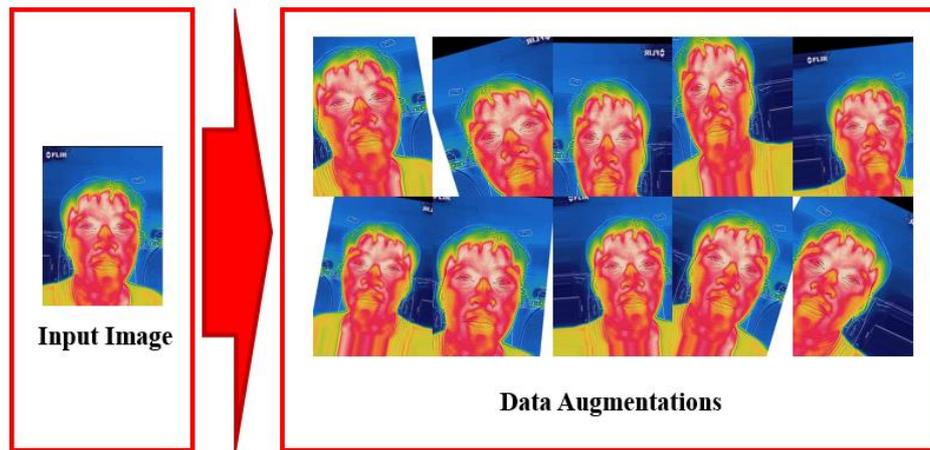


Fig. 5. Image augmentation result



Fig. 6. The face models

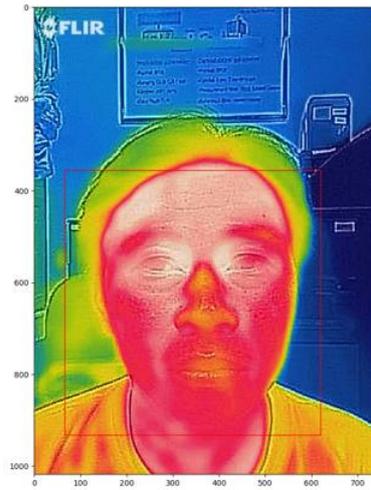


Fig. 7. Face detection of single images

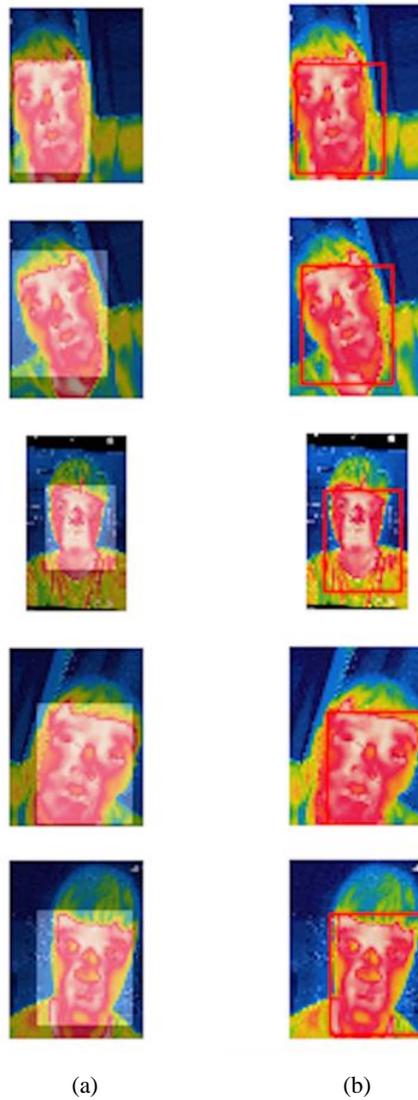


Fig. 8. Face detection of new images; (a) actual images and (b) predicted images

Beside a single image, the model was also verified by using test images. Test images mean a number of images which were prepared to test the model. The models were tested with test images which have five images. Figure 8 shows the result of image prediction. The result displayed two outputs which were actual and predicted. The predicted images visualized by square with white color. Based on the 5th model, all faces from thermal images were predicted correctly by displaying white square.

IV. Conclusion

This research proposed a segmentation method for face modelling by using thermal images. The model was created by using a Mask RCNN methods. The data collection was done by using Flir Lepton 3.5 thermal camera which is military standard camera. The model was tested by using test images which have been prepared during data preparation. A final model successfully located faces in thermal images which have contrast type. The model was successfully predicted all tested images through some experiment.

For future work, this model will be deployed in Nvidia embedded device such as Jetson Nano. Our goal is to make a portable device to measure temperature from all detected faces in frame. We will extend the dataset by re-capturing images from public area such as airport.

Declarations

Author contribution

All authors contributed equally as the main contributor of this paper. All authors read and approved the final paper.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interest

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

References

- [1] W. Sun, Y. Song, H. Zhao and Z. Jin, "A Face Spoofing Detection Method Based on Domain Adaptation and Lossless Size Adaptation," in *IEEE Access*, vol. 8, pp. 66553-66563, 2020.
- [2] F. Alqahtani, J. Banks, V. Chandran and J. Zhang, "3D Face Tracking Using Stereo Cameras: A Review," in *IEEE Access*, vol. 8, pp. 94373-94393, 2020.
- [3] R. He, X. Wu, Z. Sun and T. Tan, "Wasserstein CNN: Learning Invariant Features for NIR-VIS Face Recognition," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 7, pp. 1761-1773, 1 July 2019.
- [4] H. Hendrick, "The Halal Logo Classification by Using NVIDIA DIGITS," in *International Conference on Applied Information Technology and Innovation (ICAITI)*, 2018, pp. 162–165.
- [5] G. Scebba, G. Da Poian and W. Karlen, "Multispectral Video Fusion for Non-Contact Monitoring of Respiratory Rate and Apnea," in *IEEE Transactions on Biomedical Engineering*, vol. 68, no. 1, pp. 350-359, Jan. 2021.
- [6] C. Massaroni, D. S. Lopes, D. Lo Presti, E. Schena, and S. Silvestri, "Contactless monitoring of breathing patterns and respiratory rate at the pit of the neck: A single camera approach," *J. Sensors*, vol. 2018, .
- [7] Kasprzyk-Kucewicz, T., Cholewka, A., Bałamut, K. et al. "The applications of infrared thermography in surgical removal of retained teeth effects assessment". *J Therm Anal Calorim* , vol 144, no 1, pp 139-144, 2020.
- [8] Z. Rustam, S. Hartini, R. Y. Pratama, R. E. Yunus, and R. Hidayat, "Analysis of architecture combining Convolutional Neural Network (CNN) and kernel K-means clustering for lung cancer diagnosis," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 10, no. 3, pp. 1200–1206, 2020.
- [9] H. Almubarak, Y. Bazi, and N. Alajlan, "Two-stage mask-RCNN approach for detecting and segmenting the optic nerve head, optic disc, and optic cup in fundus images," *Appl. Sci.*, vol. 10, no. 11, 2020.
- [10] C. Zhang, X. Xu, and D. Tu, "Face Detection Using Improved Faster RCNN," no. February 2018, 2018.
- [11] L. Hao and F. Jiang, "A New Facial Detection Model based on the Faster R-CNN," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 439, no. 3, 2018.
- [12] C. Fujii, "Thermal Camera," *J. Inst. Telev. Eng. Japan*, vol. 29, no. 9, pp. 705–713, 1975.
- [13] Z. Pei, H. Xu, Y. Zhang, M. Guo, and Y. Yee-Hong, "Face recognition via deep learning using data augmentation based on orthogonal experiments," *Electron.*, vol. 8, no. 10, pp. 1–16, 2019.

- [14] K. He, G. Gkioxari, P. Dollár, and R. Girshick, “Mask R-CNN,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 2, pp. 386–397, 2020.
- [15] H. Hendrick, W. Zhi-Hao, C. Hsien-I, C. Pei-Lun, and J. Gwo-Jia, “IOS mobile APP for tuberculosis detection based on chest X-ray image,” *Proc. ICAITI 2019 - 2nd Int. Conf. Appl. Inf. Technol. Innov. Explor. Futur. Technol. Appl. Inf. Technol. Innov.*, pp. 122–125, 2019.