

# Video-based Respiratory Monitoring System for Inactive Non-human Primates

Hongru Jiang

School of Biomedical Engineering  
Shanghai Jiao Tong University  
Shanghai, China, 200240  
Email: hongrujiang@sjtu.edu.cn

XiangDong Bu

School of Biomedical Engineering  
Shanghai Jiao Tong University  
Shanghai, China, 200240  
Email: bxd0324@sjtu.edu.cn

Zhen Xu

Bio-X Institutes  
Shanghai Jiao Tong University  
Shanghai, China, 200240  
Email: xuzhen24@sjtu.edu.cn

Qianqian Shi

Department of Precision Instrument  
Tsinghua University  
Beijing, China, 100084  
Email: qqshi@mail.tsinghua.edu.cn

Huajin Tang

College of Computer Science and Technology  
Zhejiang University  
Hangzhou, China, 310027  
Email: htang@zju.edu.cn

Yao Chen

School of Biomedical Engineering  
Shanghai Jiao Tong University  
Shanghai, China, 200240  
Email: yao.chen@sjtu.edu.cn

**Abstract**—Non-human primates (NHPs) are critical in neuroscience research, where researchers implant electrodes into their brains to observe neuronal activity. However, the high post-operative mortality rate of NHPs remains a significant concern, primarily due to inadequate post-surgical monitoring. Thus, it is essential to develop a monitoring system for assessing their physiological status. Conventional wearable monitoring devices are often removed by NHPs, making it challenging to monitor their well-being. To address this issue, we propose a video-based monitoring system that utilizes Eulerian video magnification techniques and principal component analysis to detect and estimate the respiratory rate remotely. The results show that our method has a difference of less than 1 breathing rate per minute compared to a commercial monitor and works well in experiments involving different species, demonstrating the accuracy and effectiveness of our method. We found that Gaussian pyramid representation has lower requirements for the ROI selection, making it more robust for practical applications. The proposed system has the potential to improve animal welfare and facilitate new experiments in neuroscience research.

## I. INTRODUCTION

Non-human primates (NHPs), such as rhesus monkeys and marmosets, play a crucial role in neuroscience research, enabling scientists to record neuronal activity and understand the complex neural mechanisms underlying various brain functions, such as vision [1], movement [2], perception [3], decision-making [4], and memory [5]. These studies often involve the implantation of electrodes in a well-trained monkey's brain. However, the high post-operative mortality rate of these animals remains a significant challenge, largely attributed to inadequate post-surgical monitoring.

There is an urgent need for monitoring vital signals that can warn researchers timely in case of emergencies. Kunitatsu et al. [6] proposed a method to monitor the respiratory rate of NHPs by measuring the nasal air temperature using a nasal thermosensor. However, the application scenarios of wearable devices [7] are limited (Fig. 1A). NHPs are highly active and naughty animals, who often removed or damaged wear-

able devices. On the other hand, non-contact methods have overcome this limitation. These methods often rely on remote sensors, such as radar, microphone, and camera [8], to gather data physiology signals, without constraints and discomfort on subjects. Zhang et al. [9] employed a frequency-modulated continuous wave radar to monitor heart rate and respiration rate of monkeys. However, radar systems are usually more expensive than camera systems, and their installation and maintenance may be more complex.

Camera-based approaches are often simple, widely available and low-cost, amplifying the small motions in a video sequence (Fig. 1B). The motion magnification techniques can be roughly divided into Lagrangian, Eulerian approaches. Lagrangian methods are based on tracking the optical flow and uses it to move the pixels [10]. Eulerian video magnification (EVM), on the other hand, decomposes video frames into representations by an image pyramid, filters the representations that need to be amplified, and reconstructs the amplified representations [11], [12]. EVM approaches are faster and less affected by noise compared to Lagrangian-based approaches [13], successfully applying in the field of biomedical and healthcare, including the detection of pulse [14], [11], respiratory rate [15], [16] and heart rate [17], [18]. Froesel et al. [19] implemented EVM heart rate tracking method for the anesthetized and behaving rhesus monkey, where the video recording was targeted to the face. However, it is difficult to always record the monkey's face as he may sleep at any position in the cage. Therefore, monitoring heart rate maybe not a good choice.

Respiratory rate is one of the most important vital signals that directly reflect health status of animals. Brieva et al. [20] magnifying chest movements using EVM based on the Hermite transform, and classified inhalations and exhalations frames by artificial hydrocarbon network to estimate the breathing rate. But this work has a high demand for large-scale and diverse datasets for training. Zhang et al. [15] decomposed

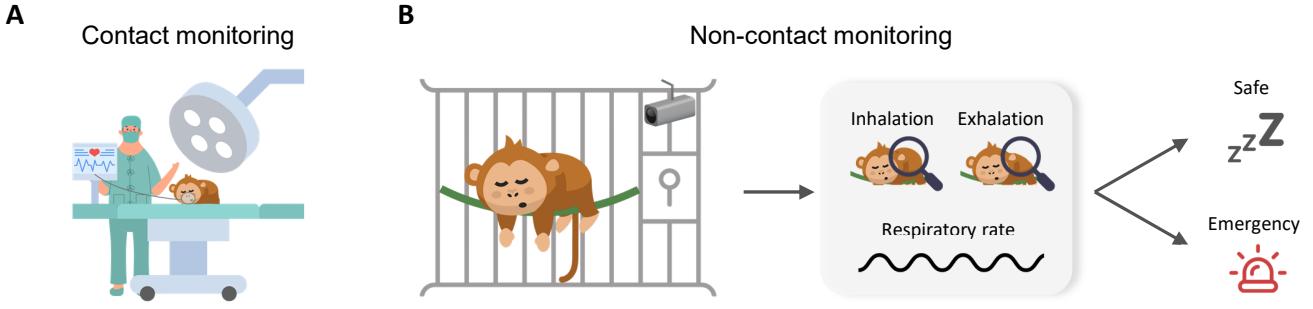


Fig. 1. Respiratory Monitoring System for Non-human Primates (A) Contact sensors for monitoring. (B) Non-contact system that magnify the subtle breathing waves from videos and evaluate the health status of a sleeping monkey.

video frames into different scales by Laplacian and Gaussian pyramids, and utilized principal component analysis (PCA) to extract breathing waveform from upper body movements. But this method ignored the selection of region of interest (ROI), which could be affected by background noise and unwanted large motions. Alam et al. [21] proposed a solution to decrease the sensitivity to unwanted movements. They reduce distortion of the salient features of the video background and utilized phase-based EVM to extract breathing motions. Mattioli et al. [16], on the other hand, used amplitude- and phase-based EVM to apply motion signals and proposed an automatic ROI selection algorithm based on the Maximum Likelihood principle to estimate the respiratory rate. But the phase-based EVM is time-consuming, and may not suitable for real-time monitoring in application.

Here, we propose a non-contact method for extracting respiratory signals from video recordings of NHPs. Prior knowledge-based and deep learning-based ROI selections for different scenarios are also proposed. Our approach combines EVM for the detection of subtle movements, and PCA for respiratory signal extraction. We evaluate the effectiveness and accuracy of our method in a experiment anesthetized rhesus monkey in the surgical setting. The results showed that the algorithm only had a difference of less than 1 breathing rate per minute compared to the commercial veterinary monitor even measured from different views. Moreover, we also applied the system in monitoring sleeping rhesus monkey and anesthetized marmoset. We compare the performance of Gaussian and Laplacian pyramids for spatial decomposition in the EVM algorithm and found that Gaussian pyramid representation has lower requirements for the ROI selection. In summary, our method offers a promising tool for monitoring the health status of NHPs in situations where wearable devices are impractical or invasive, improving animal welfare and facilitating novel experimental designs.

## II. METHOD

In this section, we first introduce the selection of ROI. Then, we explain the EVM to extract subtle breathing motions. Finally, we describe the method to estimate the respiratory rate. All the frames of the recorded videos were convert to

grayscale images when using Gaussian pyramids. The whole system is depicted in Fig. 2.

### A. Region of interest selection

To automatically select the ROI, we used the YOLO to detect monkeys in the video frames. YOLO (You Only Look Once) is a state-of-the-art, real-time object detection system that uses convolutional neural networks to predict bounding boxes and class probabilities directly from full images in one evaluation. We used the pre-trained YOLOv8 model [22] and trained it on our custom dataset of annotated NHPs images (rhesus monkey body and marmoset face) to improve its detection accuracy. Then, the fine-tuned YOLOv8 model was applied to each video frame, generating bounding boxes around the detected monkeys. The regions within these bounding boxes were then extracted from the frames, creating a new set of video sequences.

Additionally, we also employed a manual ROI selection approach to focus on specific anatomical regions of NHPs such as the abdomen, chest, and nostrils that are known to exhibit prominent respiratory motions in the video frames.

### B. Eulerian video magnification

To enhance the subtle respiratory motions of the monkeys, we applied the EVM approach [11], efficiently making breathing motions more obvious to analyze. This approach comprise two main steps: spatial decomposition and temporal filtering. First, the image is decomposed into different image representation at four scales by Gaussian pyramids. Specifically, the image is convolved with a Gaussian kernel and downsampled in both the horizontal and vertical directions:

$$G_k(i, j) = \sum_{m=-2}^2 \sum_{n=-2}^2 w(m, n) \cdot G_{k-1}(2i + m, 2j + n) \quad (1)$$

where  $G_k(i, j)$  is the pixel value at position  $(i, j)$  in the  $k$ -th level of the Gaussian pyramid,  $w(m, n)$  is the Gaussian kernel. The resulting image is upsampled by interpolation:

$$\hat{G}_k(i, j) = 4 \sum_{m=-2}^2 \sum_{n=-2}^2 w(m, n) \cdot G_{k-1}(\lfloor \frac{i}{2} \rfloor + m, \lfloor \frac{j}{2} \rfloor + n) \quad (2)$$

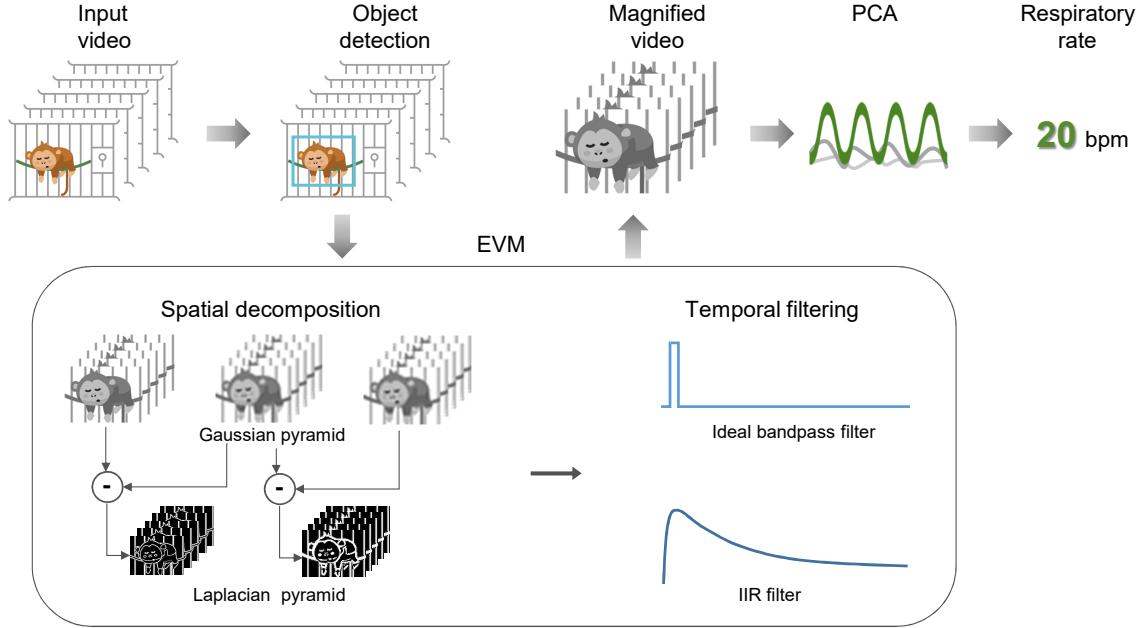


Fig. 2. Block diagram of the respiratory monitoring system

where  $\lfloor \cdot \rfloor$  denotes the floor operation.

The Laplacian pyramid is derived by subtracting each level of the Gaussian pyramid from the next lower level:

$$L_k(i, j) = G_{k-1}(i, j) - \hat{G}_k(i, j) \quad (3)$$

where  $G_0$  is the original image.

Given an image sequence  $I(x, t)$ , where  $x$  represents the pixel position and  $t$  represents time. The intensities of the pixel  $x$  can be expressed as:

$$I(x, t) = f(x + \delta(t)) \quad (4)$$

where  $f(x) = I(x, 0)$  is the first image of the video and  $\delta(t)$  is the displacements within the image.

The image can be approximated by the first order Taylor expression:

$$I(x, t) \approx f(x) + \delta(t) \frac{\partial f(x)}{\partial x} \quad (5)$$

In the temporal filtering, each representation is pixel-wise filtered to select specific frequencies of interest:

$$B(x, t) = \delta(t) \frac{\partial \hat{f}(x)}{\partial x} \quad (6)$$

For the Gaussian pyramid representation, we used the ideal band filter. For the Laplacian pyramid representation, we used the Butterworth filter of the second order with Infinite Impulse Response (IIR). The range of frequency is between 0.2-0.4 Hz to magnify breathing motions. The selected part is magnified by a factor  $\alpha$  and added back to the original image.

$$\tilde{L}(X, t) = L(X, t) + \alpha B(X, t) \quad (7)$$

where  $\tilde{L}(X, t)$  is the magnified image.

### C. Respiratory signal extraction

To extract the respiratory signal from the motion-magnified video, we applied Principal Component Analysis (PCA). PCA can capture the most significant variations in the data. For the Gaussian pyramid representation, the preserved pixies in each row are averaged. For the Laplacian pyramid representation, the RGB channel was averaged and all the pixels within the ROI were reshaped into a one-dimensional vector. PCA was applied to the time series, and the first principal component was selected as the respiratory signal. We compute the difference between peaks of respiratory signal to obtain the respiratory rate.

## III. EXPERIMENT AND DISCUSSION

We conducted three experiments to validate the effectiveness of our non-contact respiratory monitoring in different scenarios across NHPs.

### A. Monitoring anesthetized rhesus monkey in surgery

In the first experiment, we demonstrate the effectiveness of our method by comparing it with a reference respiratory signal obtained from a commercial veterinary monitor (RM700, RWD). A rhesus monkey (male, 11 year-old) was placed on an operating table. Anesthesia mask was fitted over its face to deliver the isoflurane and maintain anesthesia throughout the experiment. The smartphone camera was positioned at different angles around the operating table to capture videos of the monkey from three directions: top, front and back (Fig. 3A). The camera was placed at the height of 1.2m and a distance of 1m from the operating table, and the videos were recorded simultaneously at a resolution of  $1280 \times 720$  pixels and a frame rate of 30 fps. To obtain the reference respiratory

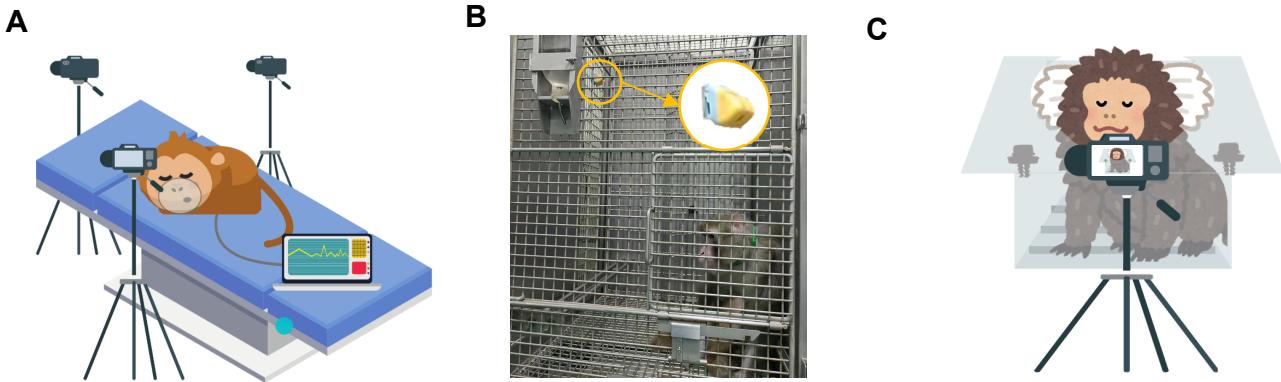


Fig. 3. Experimental Setup. (A) Anesthetized rhesus monkey in the surgical setting. (B) Rhesus monkey in cage. (C) Marmoset under mild anesthesia in primate chair.

signal, five ECG electrodes were attached to the monkey's chest and further fixed with tape.

In each view, we used two spatial decomposition (pyramids representation) approaches. From the top view, we selected the ROI of monkey's abdomen that exhibit the most obvious breathing motion (Fig. 4A, green rectangle) and used Laplacian pyramid. The EVM algorithm magnified the subtle breathing motions, making them more prominent in the video (Fig. 4A, bottom). The extracted breathing waveform exhibits clear and periodic patterns corresponding to the motion of the green line. The wires on the monkey's provide a clear mark for respiratory rate detection. Considering the absence of obvious features, we also selected another ROI that nears monkey's chest (Fig. 4A, blue rectangle). The phase of the extracted breathing waveform is nearly 1s ahead of that from the green region. Note that in some scenarios where we lack prior knowledge to select the appropriate ROI, we have to select the region that contains all the possible locations (Fig. 4A, yellow rectangle). We also used Gaussian pyramid and found it is efficient in large ROI and performs better than Laplacian pyramid. The phase of the extracted breathing waveform is almost consistent to that from the chest region (Fig. 4B). On the other hand, in the front and back views, the phase of the extracted breathing waveform in the abdomen lags behind that from the chest region by 0.3s (Fig. 4D, F). However, Gaussian pyramid representation does not work well in the back view. It detects fluctuation during 5 - 7.5 s (Fig. 4 F, bottom).

Due to the inability to obtain respiratory signals from the commercial monitor, we only compared the respiratory rate. The results showed a high agreement: Top view: Our method detected 21.3 - 21.7 breaths/min, and the veterinary monitor detected 21 breaths/min. Front view: Our method detected 20.4 - 21.6 breaths/min, and the veterinary monitor detected 20 breaths/min. Back view: Our method detected 21.1 - 23.3 breaths/min, and the veterinary monitor detected 21 breaths/min. The mean absolute error (MAE) between our method and the reference was 0.6 breaths/min, 0.5 breaths/min, and 0.9 breaths/min for the top, front and back views, respectively. These results demonstrate that our method

can achieve high accuracy in detecting respiratory signals.

#### B. Monitoring sleeping rhesus monkey at night

In the second experiment, we recorded a video of a sleeping rhesus monkey (male, 11 year-old) in its cage during the night. We used an infrared camera to capture the video, as it allowed us to record clear footage in dark conditions without disturbing the monkey's sleep. The camera was placed at the side door of the cage at the height of 1.1 m, and the video was recorded at a resolution of  $1280 \times 720$  pixels and a frame rate of 30 fps (Fig. 3 B).

We also used two pyramids representations to the recorded infrared video. The Gaussian pyramid was applied to the region that contain the monkey's body (Fig. 5 A, yellow rectangle), while the Laplacian pyramid was applied to the monkey's abdomen (Fig. 5 A, blue rectangle). The two representations display highly similar breathing waveform. The respiratory rate of sleeping monkey is 20.5 breaths/min.

#### C. Monitoring marmoset under mild anesthesia

In the third experiment, we used a marmoset (female, 3 year-old) as the subject, who was constrained in a primate chair. The monkey was under light anesthesia for the study of neural motor control. A smartphone camera was used to record a video of the marmoset's upper body at a resolution of  $1280 \times 720$  pixels and a frame rate of 30 fps. The camera was positioned at a distance of 0.8m from the primate chair (Fig. 3C).

The Gaussian pyramid was applied to the region that contain the marmoset's face (Fig. 5 C, yellow rectangle), while the Laplacian pyramid was applied to the marmoset's nose (Fig. 5 C, blue rectangle). Using different pyramids on different ROIs led to distinct results: The breathing pattern was extracted and respiratory rate of the anesthetized marmoset is 25 breaths/min (Fig. 5 D) when using Gaussian pyramid. However, our method failed to extract clear and periodic signal when using Laplacian pyramid. We also attempted to extend the scope of ROI to the marmoset's face when using Laplacian pyramid. However, the EVM magnified the micro expressions of a

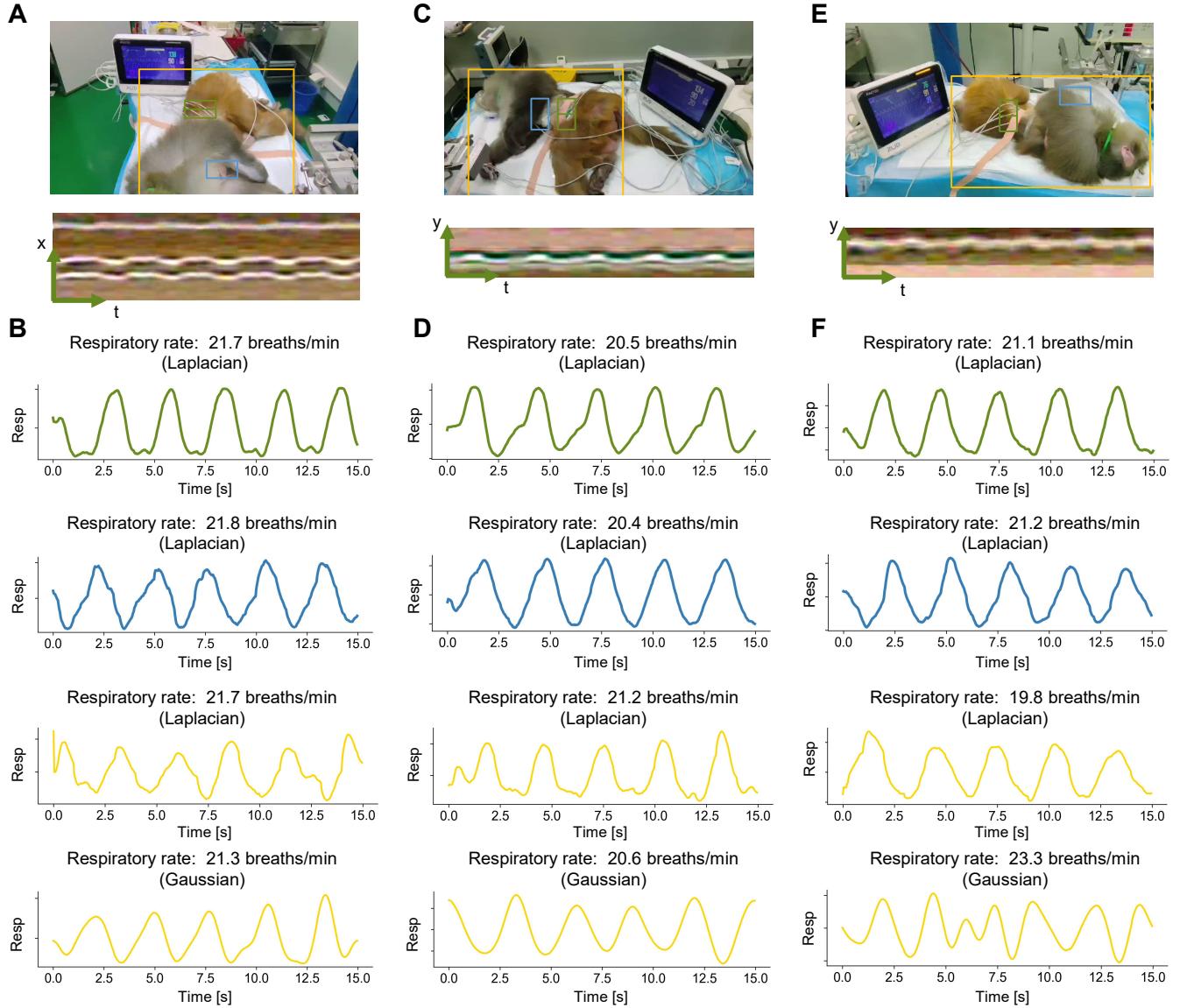


Fig. 4. Monitoring respiratory rate of the anesthetized rhesus monkey from different views. (A) Example of a video frame. Top: three ROIs are highlighted in green, blue and yellow. Bottom: a line in the green region from the magnified video plotted over time. (B) Breathing waveforms extracted from green region using Laplacian pyramid (first line), blue region using Laplacian pyramid (second line) and yellow region using Laplacian pyramid (third line) and Gaussian pyramid (fourth line). (C, E) same as (A) except for front and back views. (D, F) same as (B) except for breathing waveforms from front and back views.

marmoset monkey, such as frowning, and still failed to extract the breathing motion.

#### IV. CONCLUSION

In this study, we developed a non-contact method for monitoring respiratory rate of NHPs using a combination of deep learning, EVM and PCA. Our method was able to accurately detect the respiratory patterns of the NHPs even from different camera angles, with a mean absolute error of less than 1 breaths per minute when compared to the reference signal obtained from a commercial veterinary monitor. Moreover, We demonstrate the versatility and effectiveness of the proposed method in different experimental setups and with different monkey species. We also highlight the importance of select-

ing the appropriate spatial decomposition method and ROI for each specific experimental setup. The Gaussian pyramid appears to be more robust in large scale of ROI, while the Laplacian pyramid may be more accurate for certain ROI.

However, the linear EVM is sensitive to large motions, resulting in inaccurate respiratory monitoring in free moving animal. In the future, we will first improve the robustness of breathing detection, and improve the computation speed of video magnification algorithm for real-time applications.

#### ACKNOWLEDGMENT

This work was supported by the STI 2030—Major Projects (2022ZD0208604), and the National Natural Science Founda-

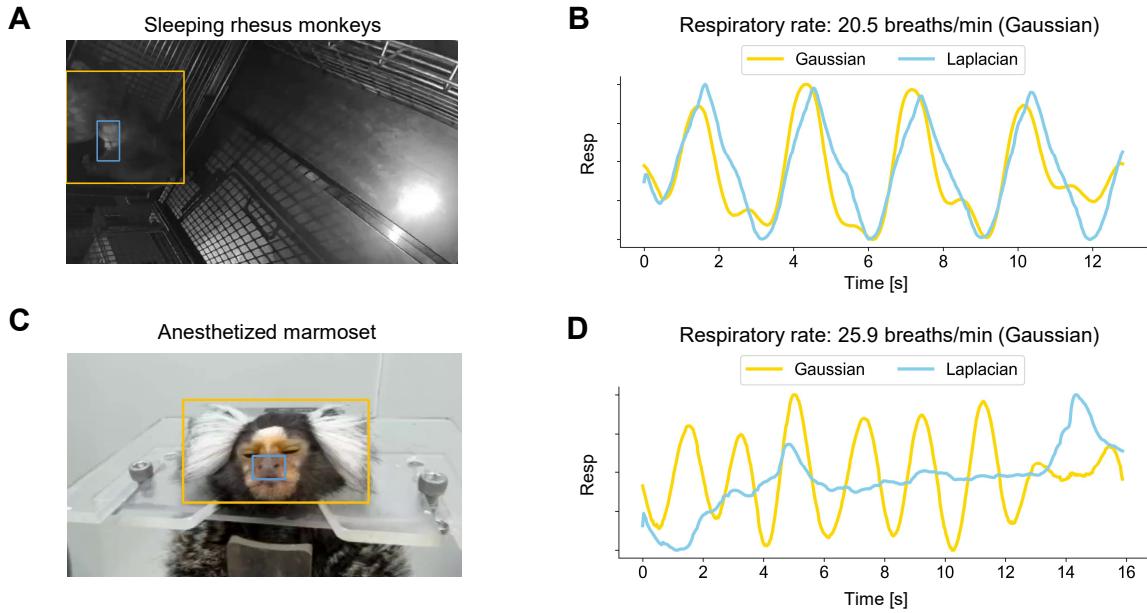


Fig. 5. Monitoring respiratory rate of NHPs. (A) Example of a video frame of sleeping rhesus monkey where two ROIs (yellow and blue) are highlighted. (B) Extracted respiratory waveforms and respiratory rate using Gaussian pyramid in yellow ROI or Laplacian pyramid in blue ROI. (C, D) Same as (A, B) except for anesthetized marmoset.

tion of China (62176151 and 61773259). We are grateful to Jieji Ren, Zhiyan Zheng and Di Zhu for discussion.

## REFERENCES

- [1] Y. Chen, S. Martinez-Conde, S. L. Macknik, Y. Bereshpolova, H. A. Swadlow, and J.-M. Alonso, “Task difficulty modulates the activity of specific neuronal populations in primary visual cortex,” *Nature neuroscience*, vol. 11, no. 8, pp. 974–982, 2008.
- [2] M. M. Churchland, J. P. Cunningham, M. T. Kaufman, J. D. Foster, P. Nuyujukian, S. I. Ryu, and K. V. Shenoy, “Neural population dynamics during reaching,” *Nature*, vol. 487, no. 7405, pp. 51–56, 2012.
- [3] A. M. Vargas, A. Bisi, A. S. Chiappa, C. Versteeg, L. E. Miller, and A. Mathis, “Task-driven neural network models predict neural dynamics of proprioception,” *Cell*, vol. 187, no. 7, pp. 1745–1761.e19, 2024.
- [4] Q. Yang, Z. Lin, W. Zhang, J. Li, X. Chen, J. Zhang, and T. Yang, “Monkey plays pac-man with compositional strategies and hierarchical decision-making,” *Elife*, vol. 11, p. e74500, 2022.
- [5] D. Mao, E. Avila, B. Caziot, J. Laurens, J. D. Dickman, and D. E. Angelaki, “Spatial modulation of hippocampal activity in freely moving macaques,” *Neuron*, vol. 109, no. 21, pp. 3521–3534, 2021.
- [6] J. Kunimatsu, Y. Akiyama, O. Toyoshima, and M. Matsumoto, “A noninvasive method for monitoring breathing patterns in nonhuman primates using a nasal thermosensor,” *eneuro*, vol. 9, no. 6, 2022.
- [7] Z. Yin, Y. Yang, C. Hu, J. Li, B. Qin, and X. Yang, “Wearable respiratory sensors for health monitoring,” *NPG Asia Materials*, vol. 16, no. 1, p. 8, 2024.
- [8] C. Massaroni, A. Nicolò, M. Sacchetti, and E. Schena, “Contactless methods for measuring respiratory rate: A review,” *IEEE Sensors Journal*, vol. 21, no. 11, pp. 12 821–12 839, 2021.
- [9] J. Zhang, R. Hu, L. Chen, Y. Gao, and D.-D. Wu, “Contactless vital signs monitoring in macaques using a mm-wave fmew radar,” *Scientific Reports*, vol. 14, no. 1, p. 13863, 2024.
- [10] N. Regev and D. Wulich, “A simple, remote, video based breathing monitor,” in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2017, pp. 1788–1791.
- [11] H.-Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. Freeman, “Eulerian video magnification for revealing subtle changes in the world,” *ACM transactions on graphics (TOG)*, vol. 31, no. 4, pp. 1–8, 2012.
- [12] N. Wadhwa, M. Rubinstein, F. Durand, and W. T. Freeman, “Phase-based video motion processing,” *ACM Transactions on Graphics (ToG)*, vol. 32, no. 4, pp. 1–10, 2013.
- [13] A. M. Ahmed, M. Abdelrazek, S. Aryal, and T. T. Nguyen, “An overview of eulerian video motion magnification methods,” *Computers & Graphics*, 2023.
- [14] X. He, R. A. Goubran, and X. P. Liu, “Using eulerian video magnification framework to measure pulse transit time,” in *2014 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*. IEEE, 2014, pp. 1–4.
- [15] Y. Zhang and F. Shang, “Noncontact extraction of breathing waveform,” in *Proceedings of the 2015 International Power, Electronics and Materials Engineering Conference*. Atlantis Press, 2015/05, pp. 782–787.
- [16] V. Mattioli, D. Alinovi, G. Ferrari, F. Pisani, and R. Raheli, “Motion magnification algorithms for video-based breathing monitoring,” *Biomedical Signal Processing and Control*, vol. 86, p. 105148, 2023.
- [17] E. Moya-Albor, J. Brieva, H. Ponce, O. Rivas-Scott, and C. Gómez-Peña, “Heart rate estimation using hermite transform video magnification and deep learning,” in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2018, pp. 2595–2598.
- [18] H. Lauridsen, S. Gonzales, D. Hedwig, K. L. Perrin, C. J. Williams, P. H. Wrege, M. F. Bertelsen, M. Pedersen, and J. T. Butcher, “Extracting physiological information in experimental biology via eulerian video magnification,” *BMC biology*, vol. 17, pp. 1–26, 2019.
- [19] M. Froesel, Q. Goudard, M. Hauser, M. Gacoin, and S. Ben Hamed, “Automated video-based heart rate tracking for the anesthetized and behaving monkey,” *Scientific Reports*, vol. 10, no. 1, p. 17940, 2020.
- [20] J. Brieva, H. Ponce, and E. Moya-Albor, “Non-contact breathing rate monitoring system using a magnification technique and convolutional networks,” in *15th international symposium on medical information processing and analysis*, vol. 11330. SPIE, 2020, pp. 181–189.
- [21] S. Alam, S. P. Singh, and U. Abeyratne, “Considerations of handheld respiratory rate estimation via a stabilized video magnification approach,” in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2017, pp. 4293–4296.
- [22] R. Varghese and M. Sambath, “Yolov8: A novel object detection algorithm with enhanced performance and robustness,” in *2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS)*. IEEE, 2024, pp. 1–6.