NC STATE UNIVERSITY

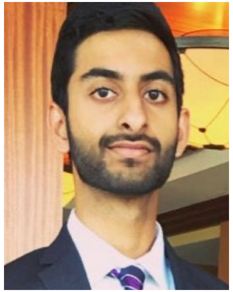
UNIVERSITY OF MARYLAND

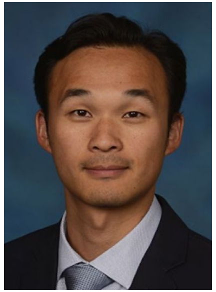
UNIVERSITY of MARYLAND SCHOOL OF MEDICINE

Remote Blood Oxygen Estimation From Video Using Neural Networks

Joshua Mathew,*[1] Xin Tian,*[2] Chau-Wai Wong,[1] Simon Ho,[3]
Donald K. Milton,[2] Min Wu[2]

[1] NC State Univ., [2] Univ. of Maryland, College Park, [3] Univ. of Maryland School of Medicine
* Equal contributions



Blood Oxygen Saturation (SpO2)

- SpO_2: Important indicator of lung functions
  - Ratio of concentration of oxygenated hemoglobin to total hemoglobin

$$SpO_2\,(\%) = \frac{C(HbO_2)}{C(Hb) + C(HbO_2)}$$

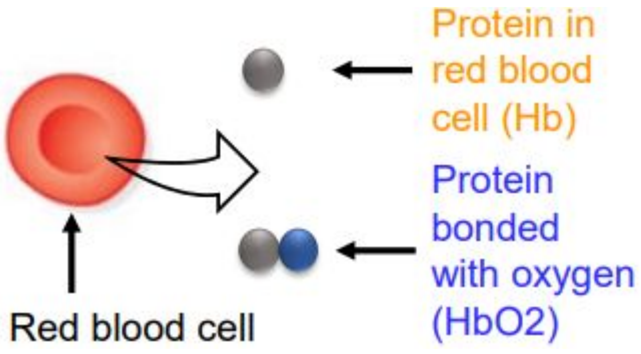
- SpO_2 Range ~95–100% in healthy people
  - COVID-19 clinical findings: Considerable # of patients had dangerously low SpO_2 w/o looking ill before conditions worsening

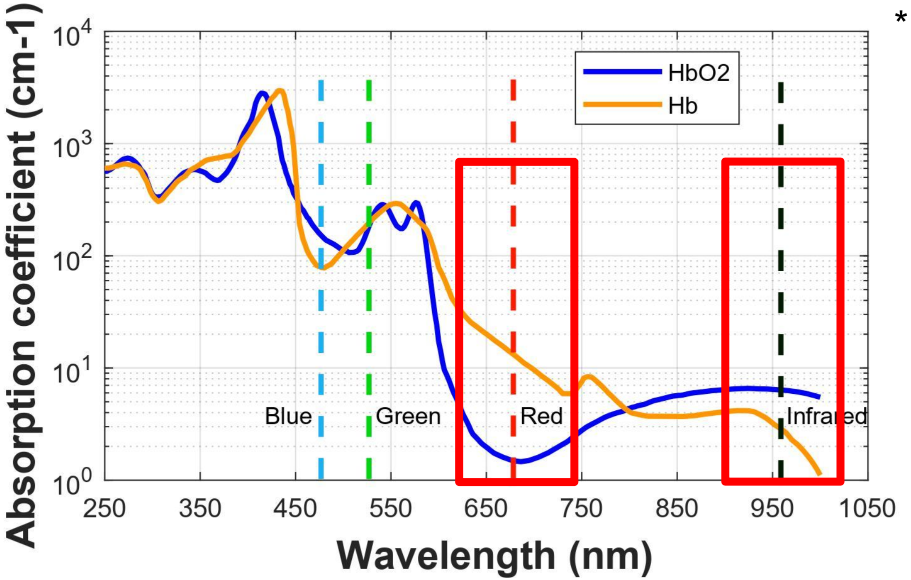
Pulse Oximeter

- Noninvasive, much less painful
- Convenient for daily use
- Less accurate (±2% error)

Pulse Oximetry

- Exploits difference in light absorption of oxygenated (HbO_2) vs deoxygenated hemoglobin (Hb) to measure blood oxygen saturation (SpO_2)





*

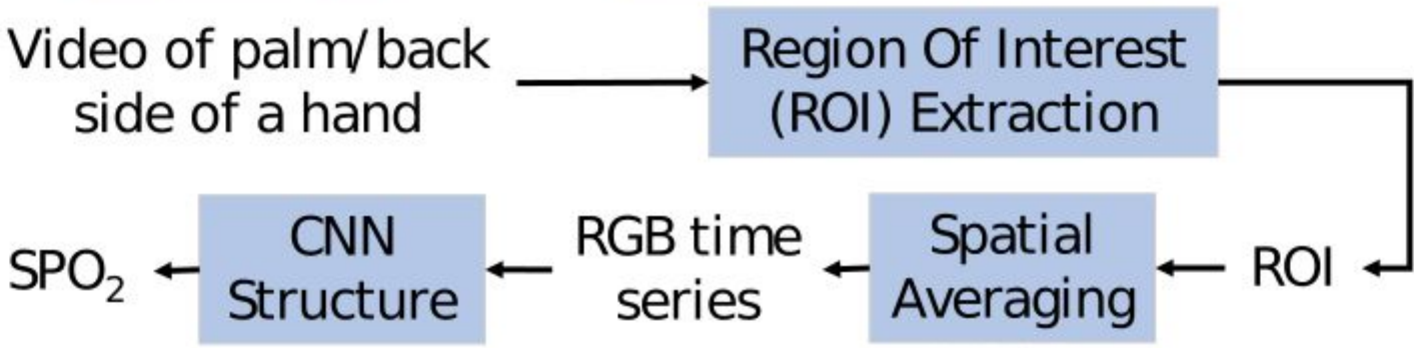
*Fig. 2 from J. Mathew, X. Tian, C.-W. Wong, S. Ho, D. K. Milton, and M. Wu, "Remote blood oxygen estimation from videos using neural networks", *IEEE Journal of Biomedical and Health Informatics*, 2023.

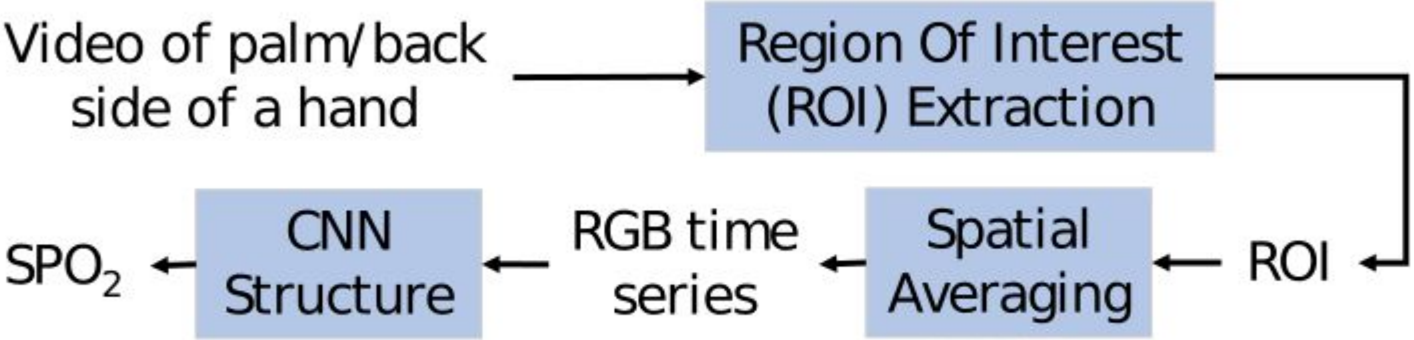
Drawbacks

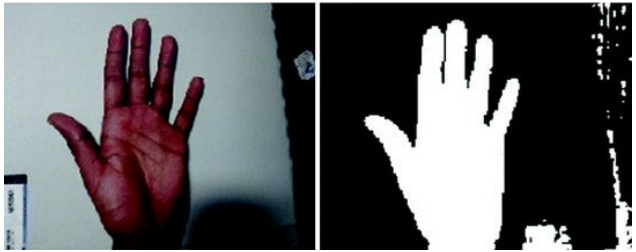
- Requires direct physical contact

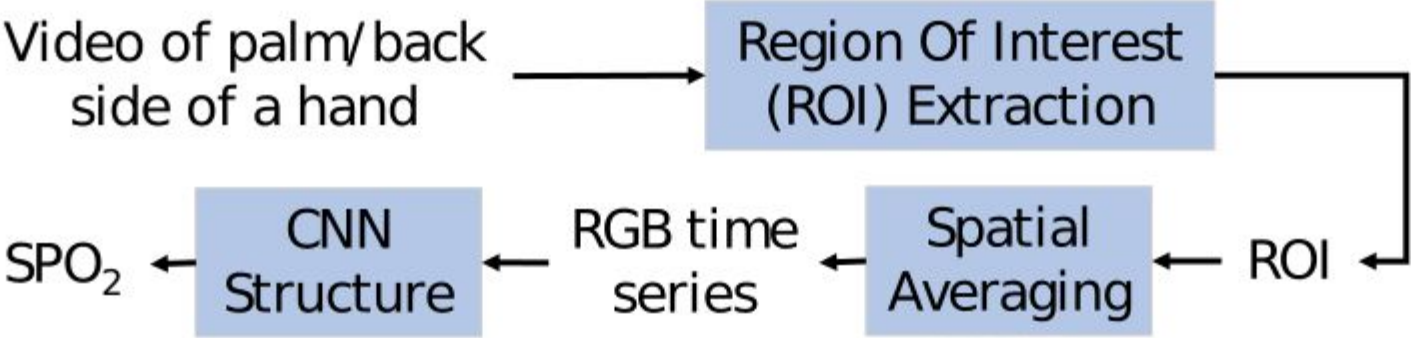
- Pulse oximeters not widely accessible, such as in marginalized communities and some undeveloped countries

Proposed Method

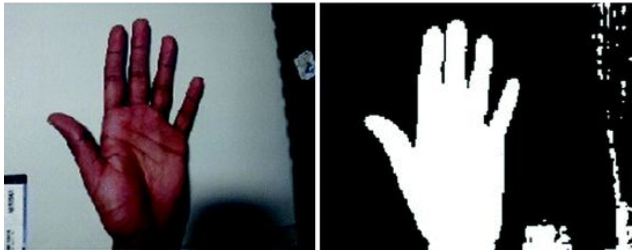


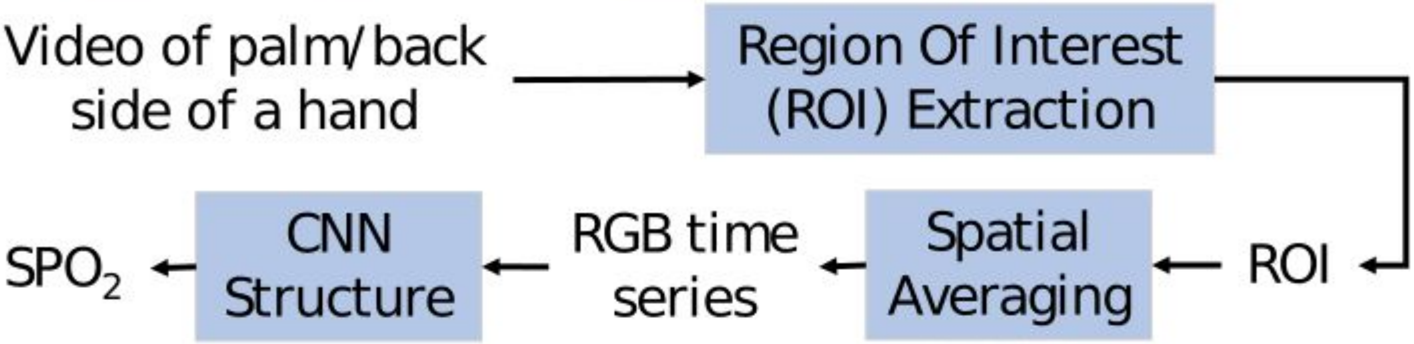

Video of palm/back side of a hand
Region Of Interest (ROI) Extraction
ROI
Spatial Averaging
RGB time series
CNN Structure
SPO_2

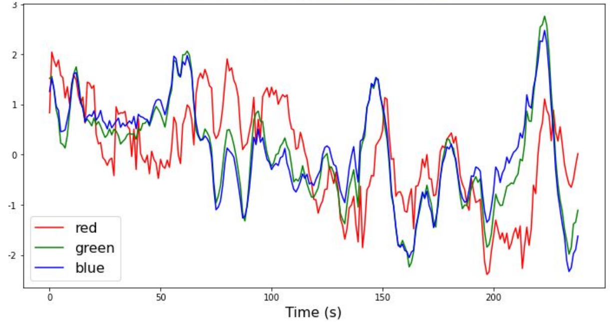


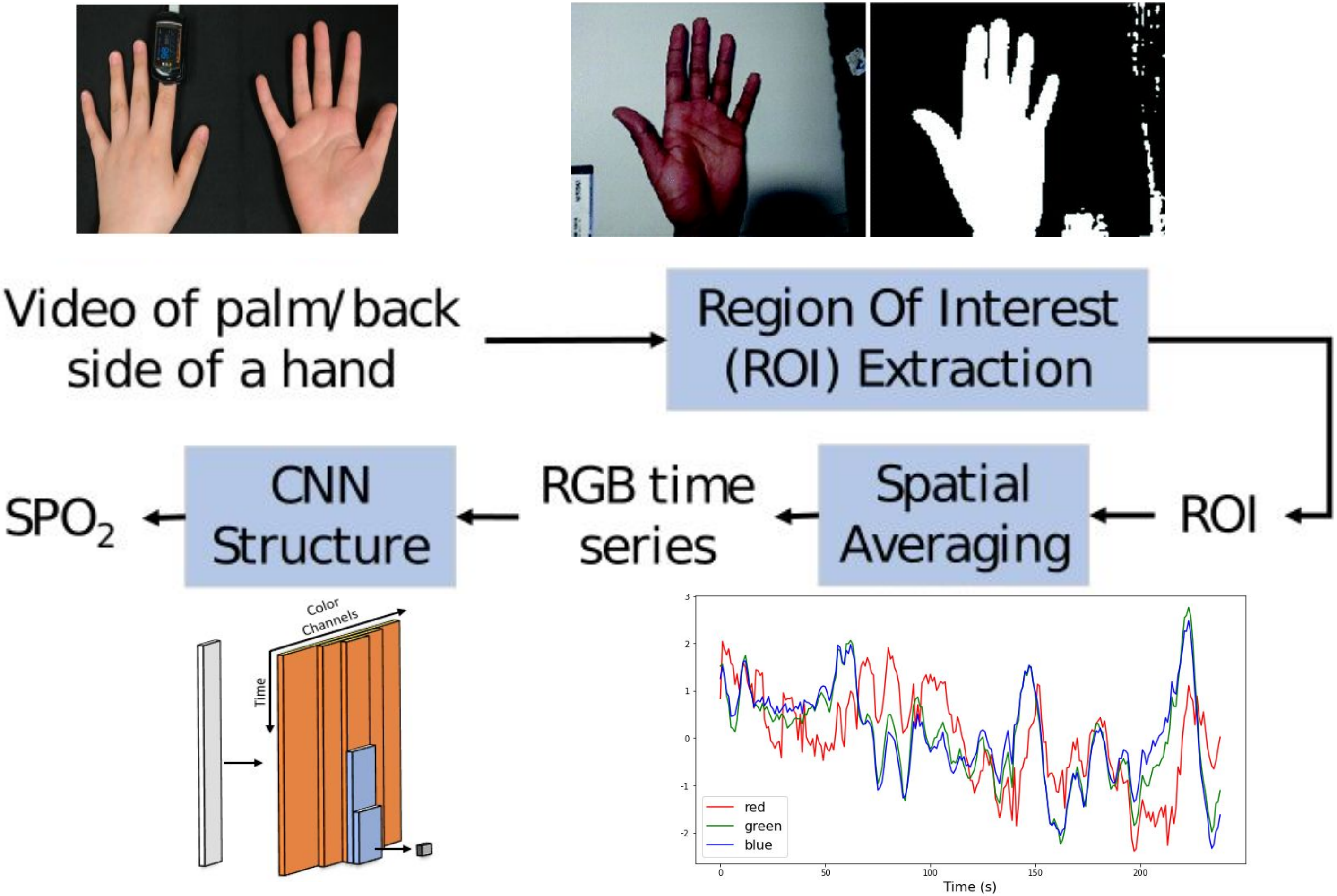


Video of palm/back side of a hand
Region Of Interest (ROI) Extraction
ROI
Spatial Averaging
RGB time series
CNN Structure
SPO_2



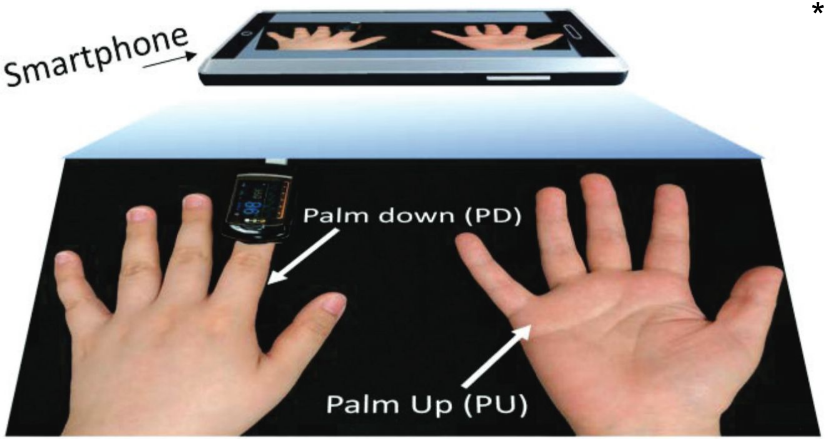

Video of palm/back side of a hand
Region Of Interest (ROI) Extraction
ROI
Spatial Averaging
RGB time series
CNN Structure
SPO_2


red
green
blue
Time (s)


Video of palm/back side of a hand
Region Of Interest (ROI) Extraction
ROI
Spatial Averaging
RGB time series
CNN Structure
SPO_2
Color Channels
Time
red
green
blue
Time (s)

Data Collection

- Data collected from videos of the hand with palm up or down



*

*Fig. 5 from J. Mathew, X. Tian, C.-W. Wong, S. Ho, D. K. Milton, and M. Wu, "Remote blood oxygen estimation from videos using neural networks", *IEEE Journal of Biomedical and Health Informatics*, 2023.

Breathing Protocol

- Participants asked to follow a breathing protocol to vary SpO_2
- Increased range of SpO_2 values

*

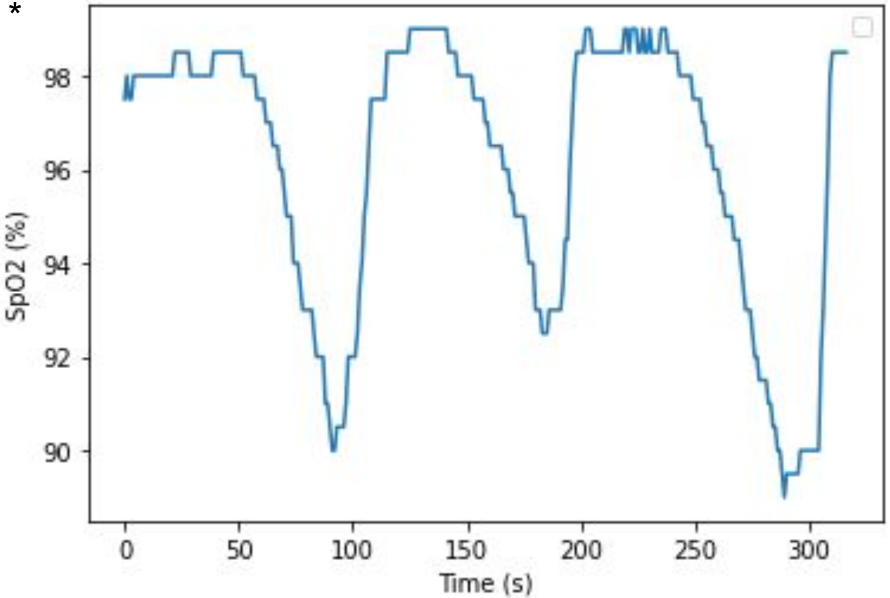
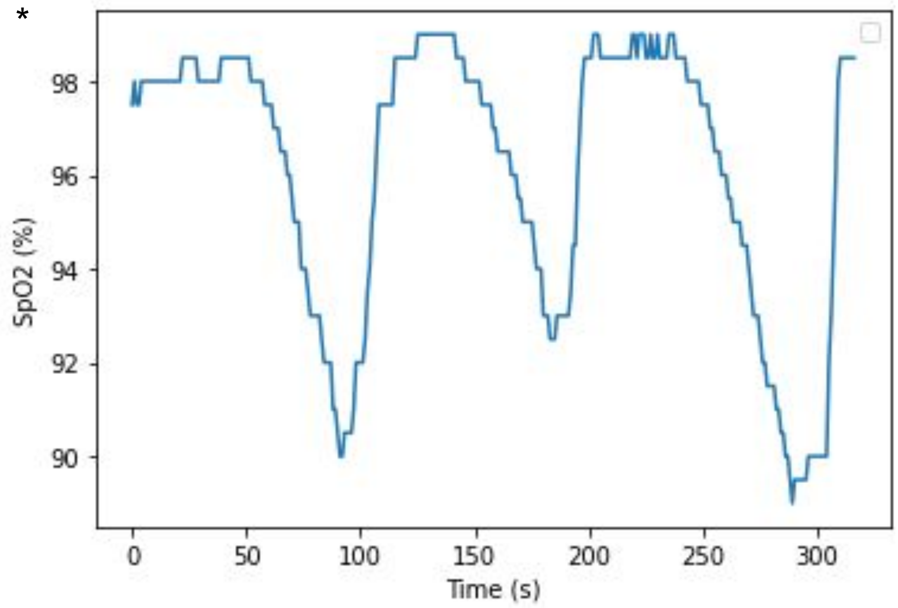


*Fig. 7a from J. Mathew, X. Tian, C.-W. Wong, S. Ho, D. K. Milton, and M. Wu, “Remote blood oxygen estimation from videos using neural networks”, *IEEE Journal of Biomedical and Health Informatics*, 2023.

Breathing Protocol

- Participants asked to follow a breathing protocol to vary SpO_2
- Increased range of SpO_2 values

*



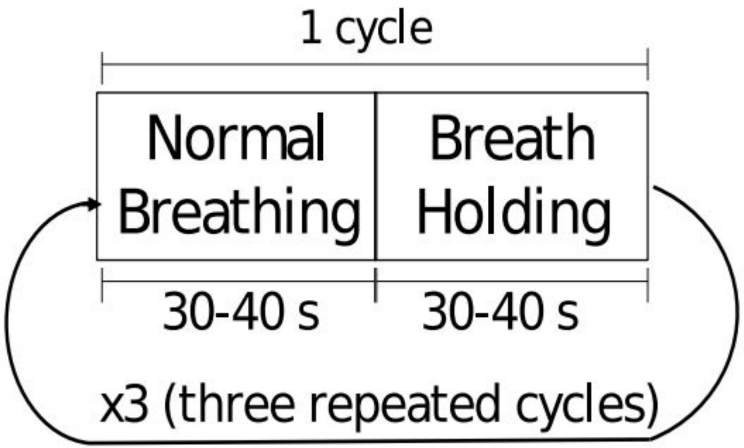


*Fig. 7a from J. Mathew, X. Tian, C.-W. Wong, S. Ho, D. K. Milton, and M. Wu, “Remote blood oxygen estimation from videos using neural networks”, *IEEE Journal of Biomedical and Health Informatics*, 2023.

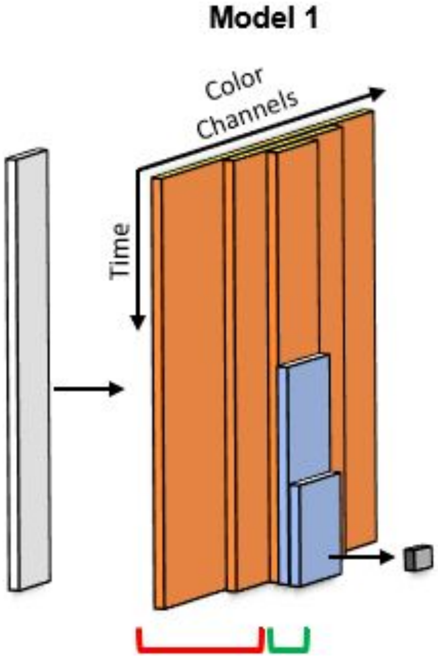
Model Structures

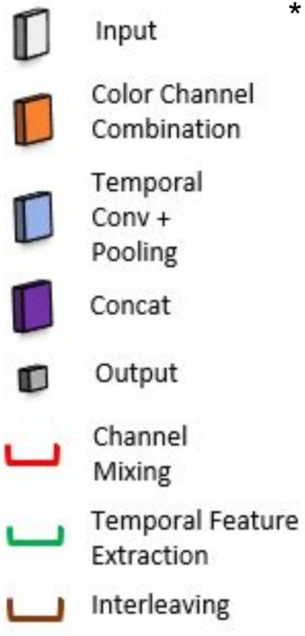
- Model 1 - Channel mixing then feature extraction
- Model 2 - Feature extraction then channel mixing
- Model 3 - Interweaving channel mixing and feature extraction

*Fig. 3 from J. Mathew, X. Tian, C.-W. Wong, S. Ho, D. K. Milton, and M. Wu, “Remote blood oxygen estimation from videos using neural networks”, *IEEE Journal of Biomedical and Health Informatics*, 2023.

Model Structures

- Model 1 - Channel mixing then feature extraction
- Model 2 - Feature extraction then channel mixing
- Model 3 - Interweaving channel mixing and feature extraction



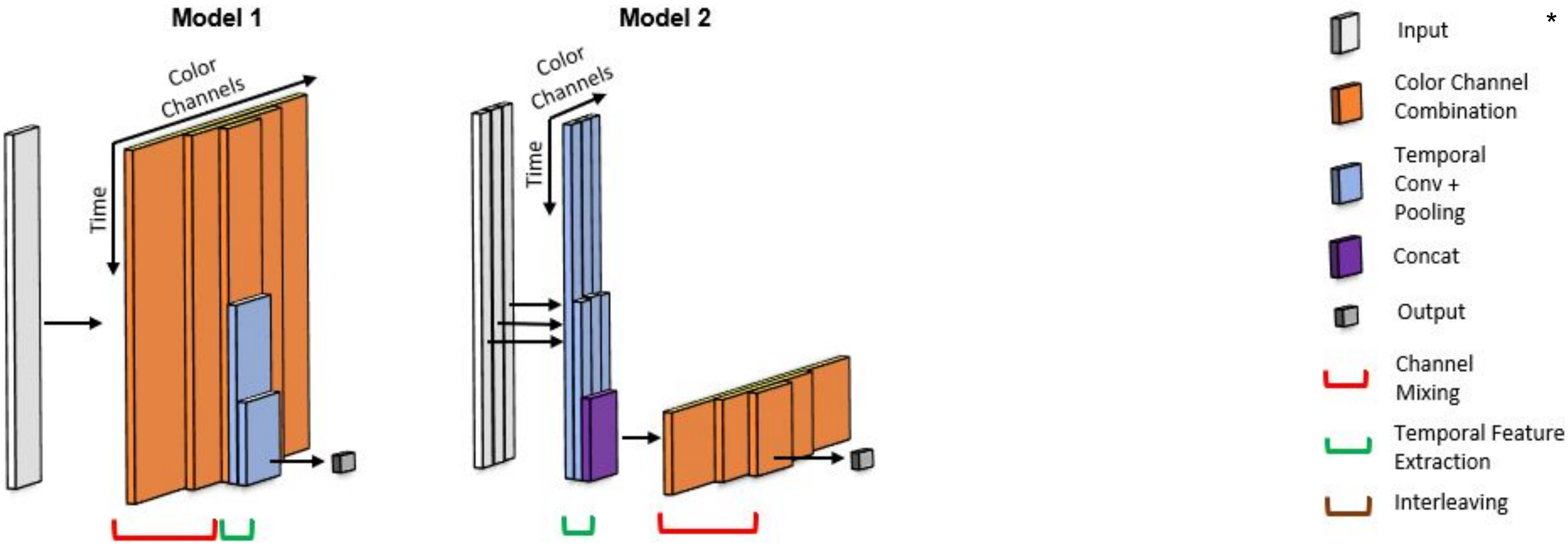


*

*Fig. 3 from J. Mathew, X. Tian, C.-W. Wong, S. Ho, D. K. Milton, and M. Wu, “Remote blood oxygen estimation from videos using neural networks”, *IEEE Journal of Biomedical and Health Informatics*, 2023.

Model Structures

- Model 1 - Channel mixing then feature extraction
- Model 2 - Feature extraction then channel mixing
- Model 3 - Interweaving channel mixing and feature extraction

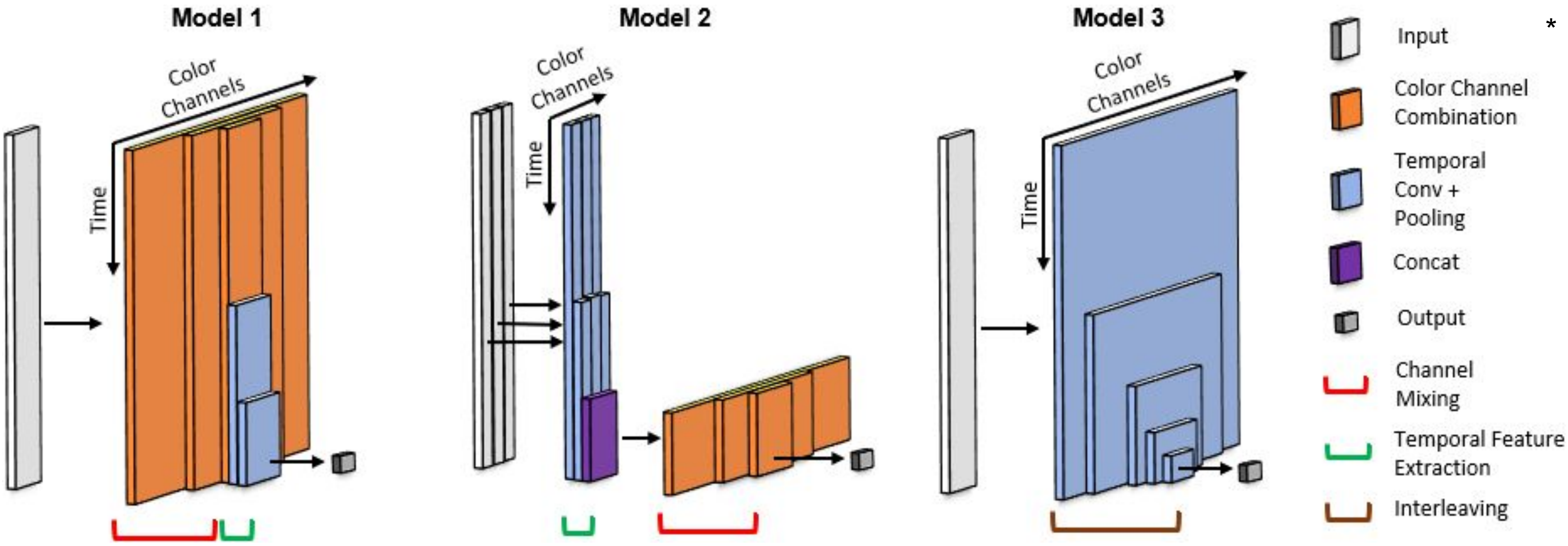


*

*Fig. 3 from J. Mathew, X. Tian, C.-W. Wong, S. Ho, D. K. Milton, and M. Wu, “Remote blood oxygen estimation from videos using neural networks”, *IEEE Journal of Biomedical and Health Informatics*, 2023.

Model Structures

- Model 1 - Channel mixing then feature extraction
- Model 2 - Feature extraction then channel mixing
- Model 3 - Interweaving channel mixing and feature extraction

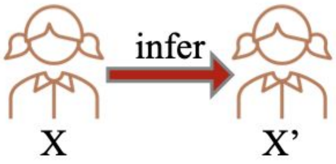


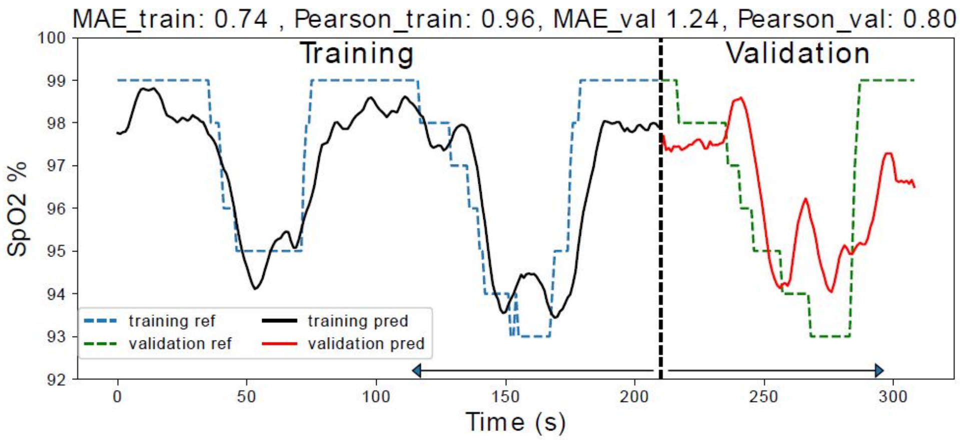
*

*Fig. 3 from J. Mathew, X. Tian, C.-W. Wong, S. Ho, D. K. Milton, and M. Wu, "Remote blood oxygen estimation from videos using neural networks", *IEEE Journal of Biomedical and Health Informatics*, 2023.

Participant Specific Testing

- 1 recording for training/validation and 1 for testing
- 2 breathing cycles for training and 1 for validation
- Train multiple instances for each model and select instance with highest validation Pearson correlation





*

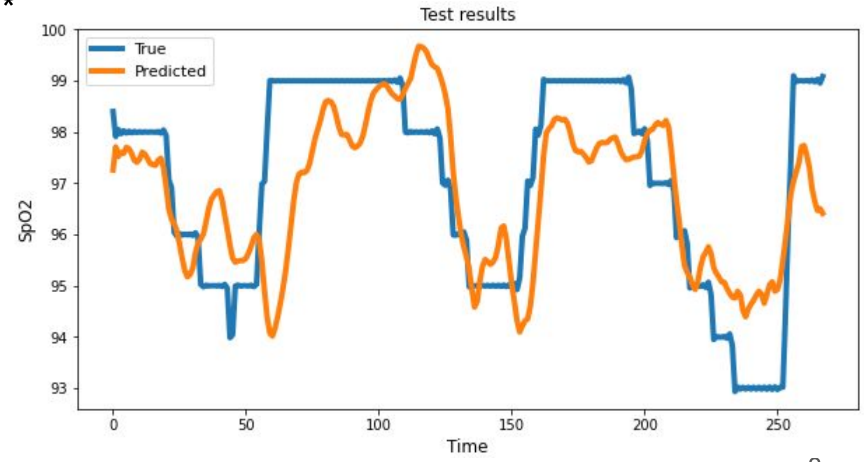
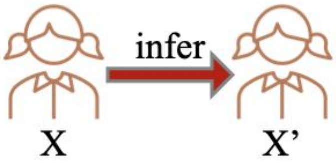
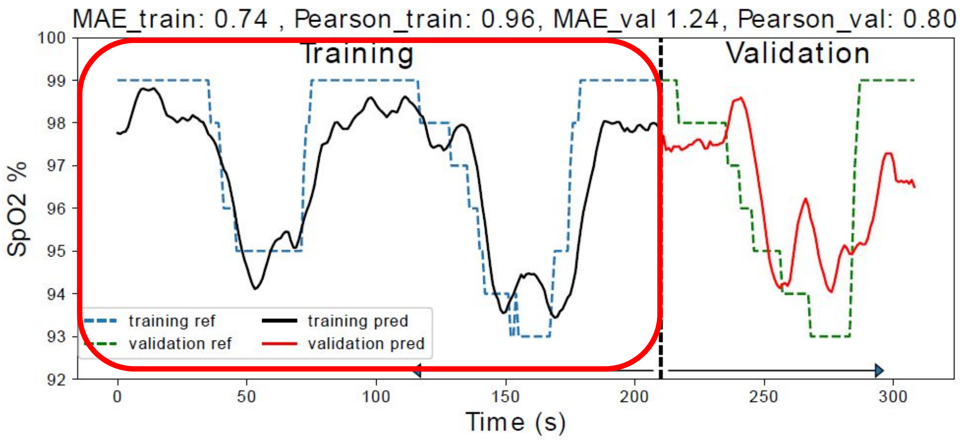


*Fig. 6b from J. Mathew, X. Tian, C.-W. Wong, S. Ho, D. K. Milton, and M. Wu, “Remote blood oxygen estimation from videos using neural networks”, *IEEE Journal of Biomedical and Health Informatics, 2023.*

Participant Specific Testing

- 1 recording for training/validation and 1 for testing
- 2 breathing cycles for training and 1 for validation
- Train multiple instances for each model and select instance with highest val pearson





*

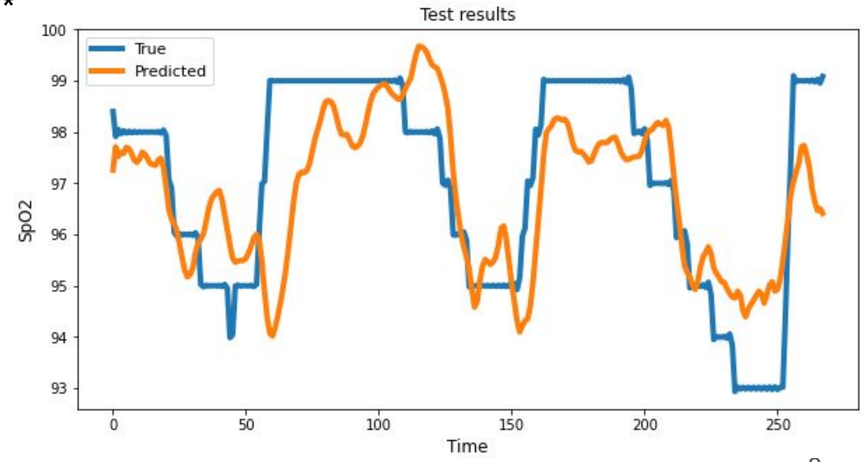
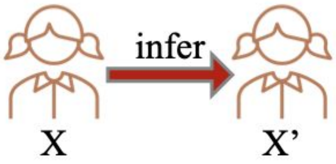
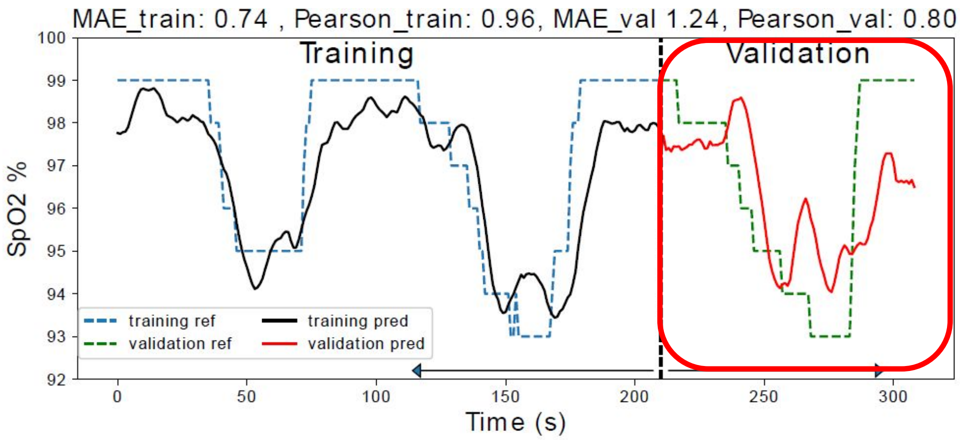


*Fig. 6b from J. Mathew, X. Tian, C.-W. Wong, S. Ho, D. K. Milton, and M. Wu, "Remote blood oxygen estimation from videos using neural networks", *IEEE Journal of Biomedical and Health Informatics, 2023.*

Participant Specific Testing

- 1 recording for training/validation and 1 for testing
- 2 breathing cycles for training and 1 for validation
- Train multiple instances for each model and select instance with highest val pearson





*

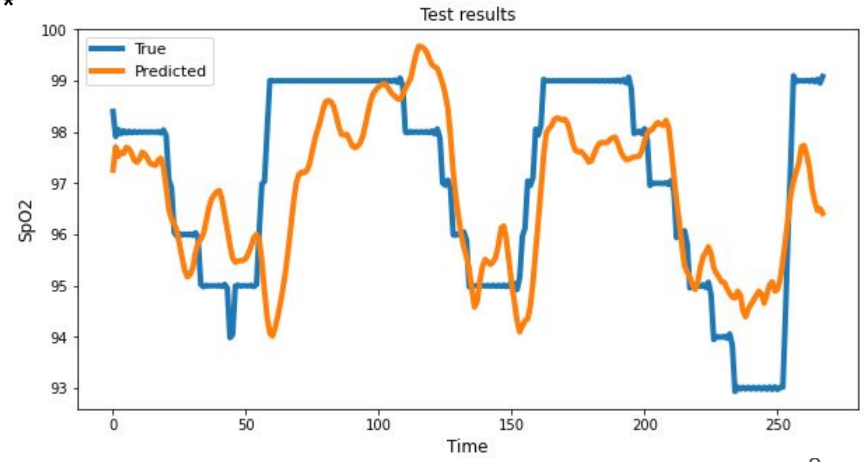
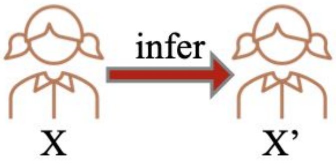
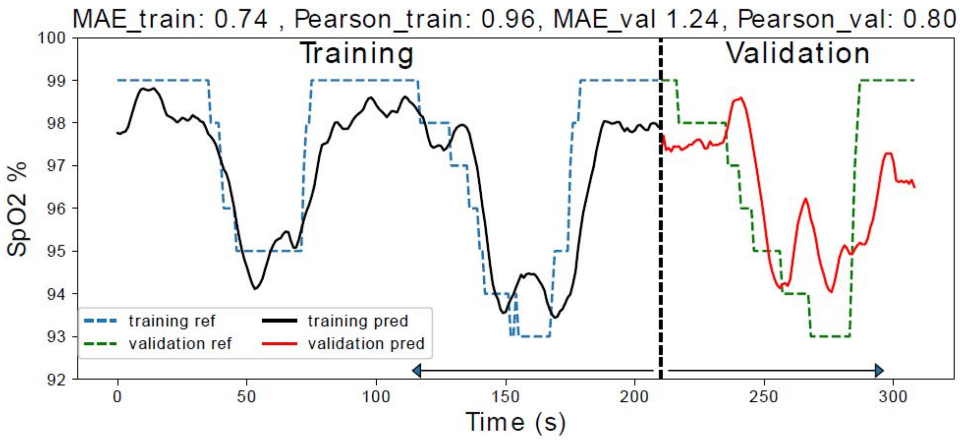


*Fig. 6b from J. Mathew, X. Tian, C.-W. Wong, S. Ho, D. K. Milton, and M. Wu, “Remote blood oxygen estimation from videos using neural networks”, *IEEE Journal of Biomedical and Health Informatics, 2023.*

Participant Specific Testing

- 1 recording for training/validation and 1 for testing
- 2 breathing cycles for training and 1 for validation
- Train multiple instances for each model and select instance with highest val pearson





*

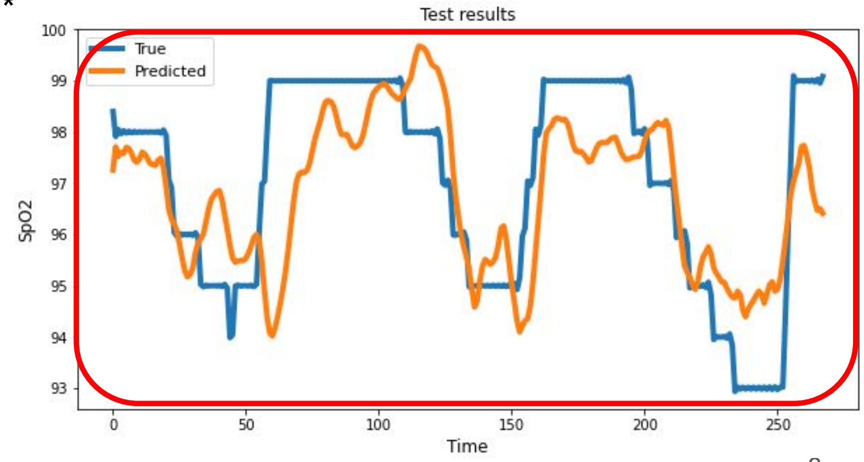


*Fig. 6b from J. Mathew, X. Tian, C.-W. Wong, S. Ho, D. K. Milton, and M. Wu, "Remote blood oxygen estimation from videos using neural networks", *IEEE Journal of Biomedical and Health Informatics, 2023.*

Participant Specific Correlation Results

- Median correlation across all participants
- Models 2 and 3 achieved best performance

	Hand Mode	Correlation Median	IQR
Model 1 (Proposed)	PD	0.41	0.40
	PU	0.39	0.37
Model 2 (Proposed)	PD	**0.46**	0.44
	PU	**0.41**	0.32
Model 3 (Proposed)	PD	0.44	0.40
	PU	**0.41**	0.46

PU (Palm Up), PD (Palm Down)

IQR (Interquartile Range) = 3rd quartile – 1st quartile

Participant Specific Correlation Results

- Ding et al. [5] – Neural network for contact-based SpO_2 prediction from video
- Scully et al. [7] – Ratio-of-ratios method

	Hand Mode	Correlation Median	Correlation IQR
Model 1 (Proposed)	PD	0.41	0.40
	PU	0.39	0.37
Model 2 (Proposed)	PD	**0.46**	0.44
	PU	**0.41**	0.32
Model 3 (Proposed)	PD	0.44	0.40
	PU	**0.41**	0.46
Scully *et al.* [5]	PD	0.08	0.37
	PU	0.19	0.24
Ding *et al.* [7]	PD	0.38	0.39
	PU	0.34	0.56

[5] X. Ding, D. Nassehi, and E. C. Larson, "Measuring oxygen saturation with smartphone cameras using convolutional neural networks," *IEEE J Biomed. Health Info.*, 2018.
[7] C. G. Scully, J. Lee, J. Meyer, et al., "Physiological parameter monitoring from optical recordings with a mobile phone," *IEEE Trans. Biomed. Eng.*, 2011.

Participant Specific MAE Results

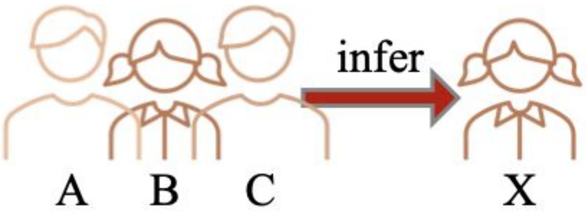
- All models outperformed Ding et al.
- Scully et al. achieves lower mean absolute error (MAE) than some models but has low correlation

	Hand Mode	Correlation Median	Correlation IQR	MAE (%) Median	MAE (%) IQR
Model 1 (Proposed)	PD	0.41	0.40	2.12	0.91
	PU	0.39	0.37	2.16	1.80
Model 2 (Proposed)	PD	**0.46**	0.44	2.09	1.32
	PU	**0.41**	0.32	1.96	0.68
Model 3 (Proposed)	PD	0.44	0.40	**1.93**	1.11
	PU	**0.41**	0.46	**1.81**	1.83
Scully *et al.* [5]	PD	0.08	0.37	1.94	0.92
	PU	0.19	0.24	2.01	0.80
Ding *et al.* [7]	PD	0.38	0.39	3.25	2.85
	PU	0.34	0.56	3.40	3.16

Naïve predictor is not useful

Leave-One-Participant-Out Experiments

- Aim: To test generalizability to unseen participants
- Train on all participant recordings but leave out test participant
- Participant-wise cross-validation



Leave-One-Out Correlation Results

- More difficult than participant-specific case
- Doing channel combination first may be more generalizable

	Hand Mode	Correlation Median	Correlation IQR
Model 1 (Proposed)	PD	**0.33**	0.42
	PU	**0.46**	0.36
Model 2 (Proposed)	PD	0.15	0.50
	PU	0.33	0.39
Model 3 (Proposed)	PD	0.23	0.38
	PU	0.27	0.31

Leave-One-Out Correlation Results

- Proposed models outperform state-of-the-art

	Hand Mode	Correlation Median	Correlation IQR
Model 1 (Proposed)	PD	**0.33**	0.42
	PU	**0.46**	0.36
Model 2 (Proposed)	PD	0.15	0.50
	PU	0.33	0.39
Model 3 (Proposed)	PD	0.23	0.38
	PU	0.27	0.31
Scully *et al.* [5]	PD	0.05	0.43
	PU	0.01	0.54
Ding *et al.* [7]	PD	0.11	0.56
	PU	0.26	0.42

Leave-One-Out MAE Results

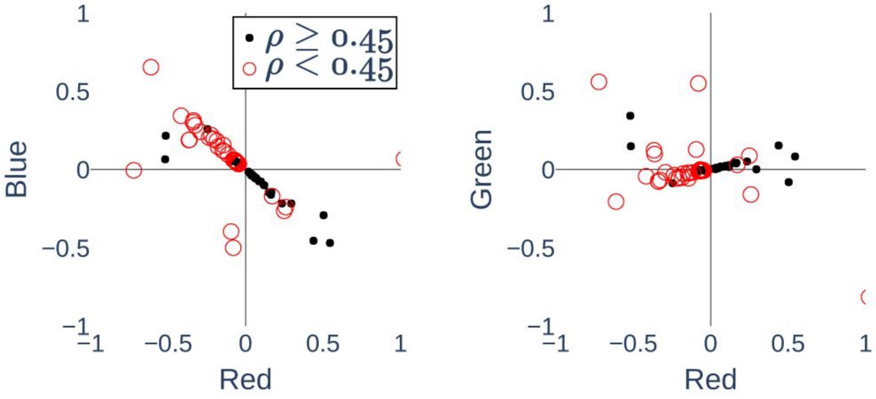
- More difficult than participant-specific case
- Features extracted after channel combination maybe more generalizable

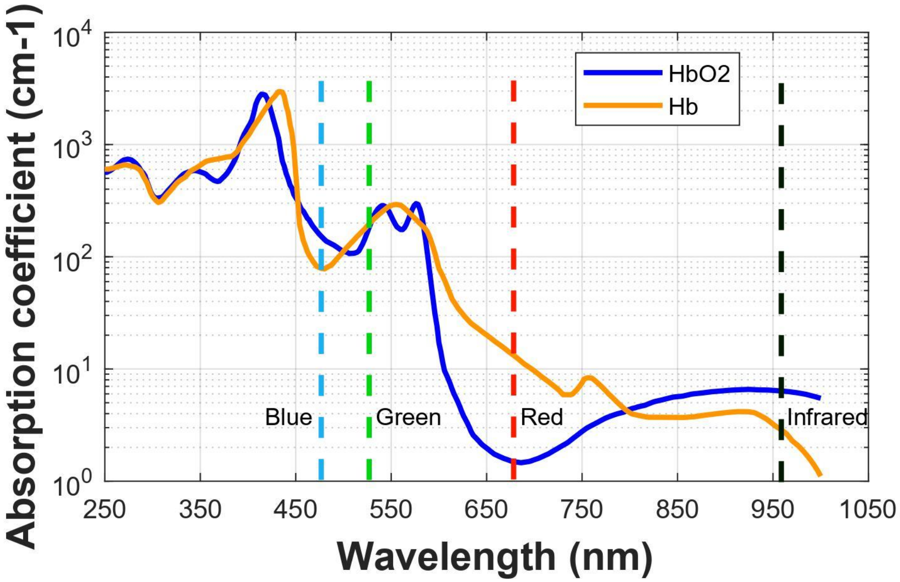
	Hand Mode	Correlation Median	Correlation IQR	MAE (%) Median	MAE (%) IQR
Model 1 (Proposed)	PD	**0.33**	0.42	2.33	1.07
	PU	**0.46**	0.36	**1.97**	0.80
Model 2 (Proposed)	PD	0.15	0.50	2.43	0.94
	PU	0.33	0.39	2.08	0.73
Model 3 (Proposed)	PD	0.23	0.38	2.48	1.18
	PU	0.27	0.31	2.02	1.03
Scully *et al.* [5]	PD	0.05	0.43	**2.08**	0.65
	PU	0.01	0.54	2.08	0.60
Ding *et al.* [7]	PD	0.11	0.56	3.19	1.61
	PU	0.26	0.42	2.43	1.22

Naïve predictor is not useful

Channel Combination Visualization

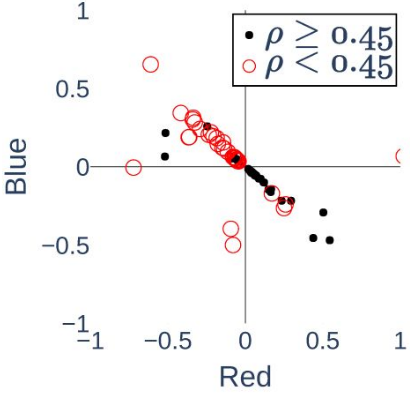
- Visualized model weights

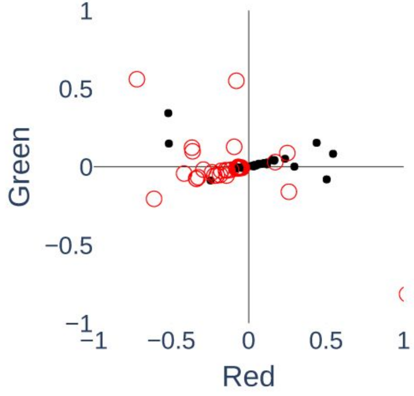


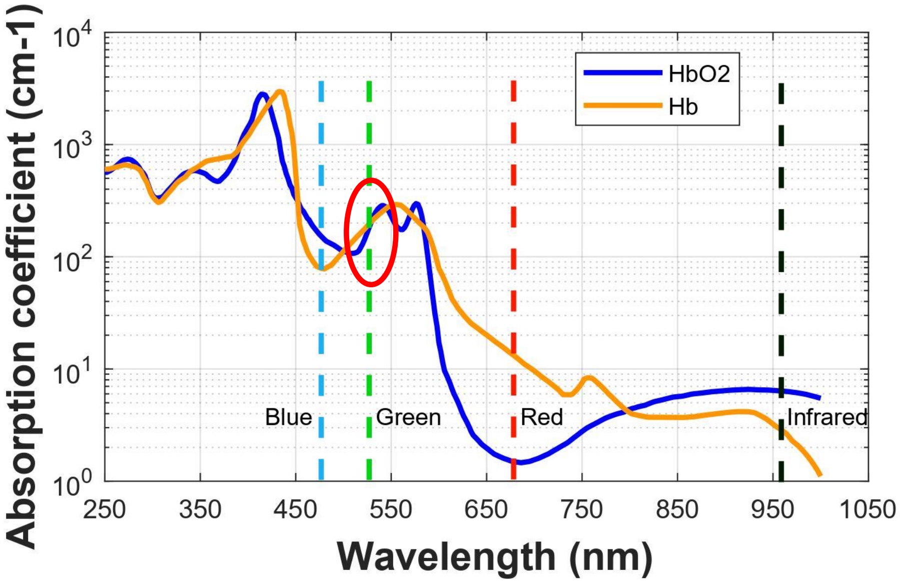


Channel Combination Visualization

- Green channel is given low weight

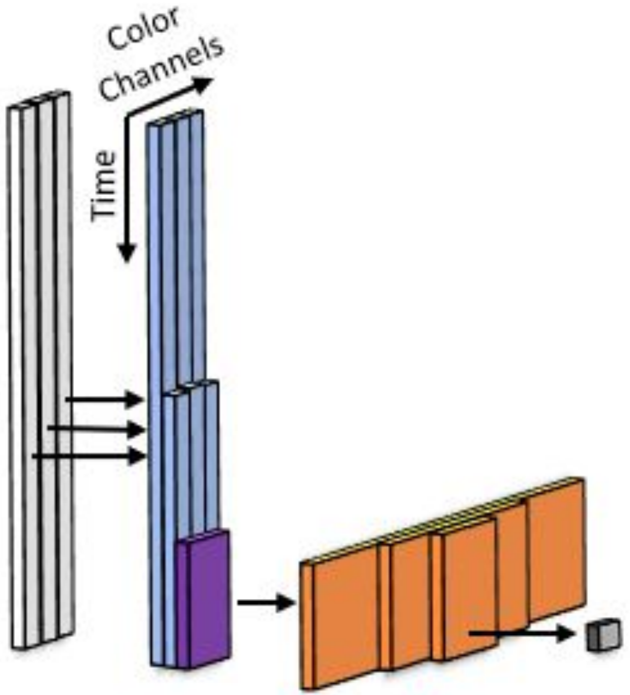






Conclusion

- Measure blood oxygen without physical contact utilizing smartphone videos
- Developeded novel neural network architectures which outperformed state-of-the-art methods
- Showed that learned RGB weights are consistent with optophysiological methods.



NC STATE UNIVERSITY
UNIVERSITY OF MARYLAND
UNIVERSITY of MARYLAND SCHOOL OF MEDICINE

Remote Blood Oxygen Estimation From Video Using Neural Networks

Joshua Mathew,*[1] Xin Tian,*[2] Chau-Wai Wong,[1] Simon Ho,[3]
Donald K. Milton,[2] Min Wu[2]

[1] NC State Univ., [2] Univ. of Maryland, College Park, [3] Univ. of Maryland School of Medicine
* Equal contributions