Transitioning sleeping position detection in late pregnancy using computer vision from controlled to real-world settings: an observational study

Allan Kember J[1], Hafsa Zia[2], Praniya Elangainesan[2], Min-En Hsieh[3], Ramak Adijeh[4], Ivan Li[5], Leah Ritchie[6], Sina Akbarian[7], Babak Taati[5], Sebastian Hobson[1], and Elham Dolatabadi[8]

[1]University of Toronto
[2]University of Toronto Temerty Faculty of Medicine
[3]National Cheng Kung University College of Electrical Engineering and Computer Science
[4]Northeastern University
[5]University of Toronto Department of Computer Science
[6]University of Toronto Scarborough Department of Biological Sciences
[7]Vector Institute
[8]University of Toronto Institute of Health Policy Management and Evaluation

August 23, 2023

Abstract

Objective: To build a computer vision model that can automatically detect sleeping position in the third trimester under real-world conditions. **Design**: This study used data from an ongoing observational study and a previous cross-sectional study. **Setting**: Participants' homes. **Sample**: Pregnant participants in the third trimester and their bed partners. **Methods**: Real-world overnight video recordings were collected from an ongoing, Canada-wide, prospective, four-night, home sleep apnea study and controlled-setting video recordings were used from a previous study. Images were extracted from the videos and body positions were annotated. Five-fold cross validation was used to train, validate, and test a model using state-of-the-art deep convolutional neural networks. **Main Outcome Measures**: Precision and recall of the model for detecting thirteen pre-defined body positions. **Results**: The dataset contained 39 pregnant participants, 13 bed partners, 12,930 images, and 47,001 annotations. The model was trained to detect pillows, twelve sleeping positions, and a sitting position in both the pregnant person and their bed partner simultaneously. The model significantly outperformed a previous similar model for the three most commonly occurring natural sleeping positions in pregnant and non-pregnant adults, with an 82-to-89% average probability of correctly detecting them and a 15-to-19% chance of failing to detect them when any one of them is present. **Conclusions**: The model holds potential to solve yet unanswered research and clinical questions regarding the relationship between sleeping position and pregnancy outcomes.

Title : Transitioning sleeping position detection in late pregnancy using computer vision from controlled to real-world settings: an observational study

Authors :

Allan J Kember, MD, MSc; Department of Obstetrics and Gynaecology, University of Toronto; Institute of Health Policy, Management, and Evaluation, University of Toronto; 123 Edward Street, Suite 1200, Toronto, Ontario, Canada, M5G 0A8.

Hafsa Zia, BSc; Temerty Faculty of Medicine, University of Toronto; Medical Sciences Building, 1 King's College Circle, Toronto, Ontario, Canada, M5S 1A8.

Praniya Elangainesan, BSc; Temerty Faculty of Medicine, University of Toronto; Medical Sciences Building, 1 King's College Circle, Toronto, Ontario, Canada, M5S 1A8.

Min-En Hsieh, MSc; Electrical Engineering and Computer Science, National Cheng Kung University; No.1, University Road, Tainan City 701, Taiwan.

Ramak Adijeh, PhD; Regulatory Affairs Program, Northeastern University; First Canadian Place, 100 King Street West, Suite 4620, Toronto, Ontario, Canada, M5X 1E2.

Ivan Li; Department of Computer Science, University of Toronto; 40 St. George Street, Toronto, Ontario, Canada, M5S 2E4.

Leah Ritchie; Department of Biological Sciences, University of Toronto Scarborough; 1265 Military Trail, Toronto, Ontario, Canada, M1C 1A4.

Sina Akbarian, MASc; Vector Institute; 661 University Avenue, Suite 710, Toronto, Ontario, Canada, M5G 1M1.

Babak Taati, PhD, PEng; KITE Research Institute, Toronto Rehabilitation Institute, University Health Network; Department of Computer Science, University of Toronto; 40 St. George Street, Toronto, Ontario, Canada M5S 2E4.

Sebastian R Hobson, MD, MPH, PhD; Department of Obstetrics and Gynaecology, University of Toronto; Maternal-Fetal Medicine Division, Department of Obstetrics and Gynaecology, Mount Sinai Hospital; Toronto, Ontario, Canada; 123 Edward Street, Suite 1200, Toronto, Ontario, Canada, M5G 0A8

Elham Dolatabadi, PhD, MESc; School of Health Policy and Management, York University, Institute of Health Policy, Management, and Evaluation, University of Toronto; 155 College Street, 4th Floor, Toronto, Ontario, Canada, M5T 3M6.

Corresponding Author :

Allan J Kember, MD, MSc; Department of Obstetrics and Gynaecology, University of Toronto; Institute of Health Policy, Management, and Evaluation, University of Toronto; 123 Edward Street, Suite 1200, Toronto, Ontario, Canada, M5G 0A8; +1-437-345-9080; allan.kember@mail.utoronto.ca

Short Title : Vision-based sleeping position detection for late pregnancy

Abstract

Objective : To build a computer vision model that can automatically detect sleeping position in the third trimester under real-world conditions.

Design : This study used data from an ongoing observational study and a previous cross-sectional study.

Setting : Participants' homes.

Sample : Pregnant participants in the third trimester and their bed partners.

Methods : Real-world overnight video recordings were collected from an ongoing, Canada-wide, prospective, four-night, home sleep apnea study and controlled-setting video recordings were used from a previous study. Images were extracted from the videos and body positions were annotated. Five-fold cross validation was used to train, validate, and test a model using state-of-the-art deep convolutional neural networks.

Main Outcome Measures : Precision and recall of the model for detecting thirteen pre-defined body positions.

Results : The dataset contained 39 pregnant participants, 13 bed partners, 12,930 images, and 47,001 annotations. The model was trained to detect pillows, twelve sleeping positions, and a sitting position in both the pregnant person and their bed partner simultaneously. The model significantly outperformed a previous similar model for the three most commonly occurring natural sleeping positions in pregnant and non-pregnant adults, with an 82-to-89% average probability of correctly detecting them and a 15-to-19% chance of failing to detect them when any one of them is present.

Conclusions : The model holds potential to solve yet unanswered research and clinical questions regarding the relationship between sleeping position and pregnancy outcomes.

Funding : Mitacs (Grant No. IT 26655)

Keywords : artificial intelligence, classification, deep learning, detection, foetal growth, localization, real-world, controlled-setting, small-for-gestational age, stillbirth, supine

Funding

This study was funded by Mitacs through a Mitacs Entrepreneur-Accelerate Program grant (No. IT 26655), which included external peer review for scientific quality but did not include priority assessment. Mitacs had no role in conducting the research and writing this paper.

Introduction

The supine going-to-sleep position, when adopted after 28 weeks of pregnancy, is associated with giving birth to a small-for-gestational-age infant and late stillbirth.[1,2] While some professional associations, including the Royal College of Obstetricians and Gynaecologists, have incorporated this evidence into clinical practice guidelines,[3] the National Institute for Health and Care Excellence has pointed out that the evidence underlying this association is based on retrospective studies of self-reported going-to-sleep position, which may be limited by inaccuracies and recall bias[4] and does not account for the variability in sleeping position following sleep onset. As such, the potential impact of sleeping position from 28 weeks through birth on pregnancy outcomes has not yet been prospectively verified with objective measurements.

We previously built a computer vision model ("SLeeP AIDePt-1"; Sleep in Late Pregnancy: Artificial Intelligence for the Detection of Position) to automatically detect sleeping positions during pregnancy from video recordings. This model was built using a video dataset captured from pregnant individuals in their third trimester simulating a range of sleeping positions in a controlled setting.[5] However, real-world generalizability and ecological validity was severely limited because the video dataset only contained a single person in the bed, used thin and pattern-free bed sheets, did not allow any objects on the bed besides head pillows, did not include prone posture, and, overall, did not capture the complexities of real-world scenarios.

In an attempt to transition to real-world settings (the present study), our objective was to expand our controlled-setting dataset to include real-world videos and develop a new model ("SLeeP AIDePt-2") for automated, unobtrusive, and non-contact detection and measurement of the sleeping position of a pregnant individual and their bed partner simultaneously overnight throughout the third trimester in the home setting. We aim to equip researchers with a tool employing this model to enable them to either confirm or disprove the associations between supine sleeping position in late pregnancy and adverse outcomes.

Methods

Design

This study leverages real-world, overnight, video recordings collected from an ongoing, prospective, observational, four-night, home sleep apnea study (ClinicalTrials.gov Identifier: NCT05376475) and controlled-setting video recordings from a cross-sectional study, which is open-access and can be found online.[5]A core outcome set was not used in the design of this study.

Participants

Participants were recruited Canada-wide by the research team via various social media platforms. All data collection was completed within the participants' own homes.

Participants for the real-world study were eligible to take part if their American Society of Anesthesiologists Physical Status (ASA PS) class was II or lower, they had a low-risk singleton pregnancy, were in the third trimester (between 28 weeks and 0 days through 40 weeks and 6 days gestation, inclusive, determined by first-trimester ultrasound), were aged 18 to 50 years, had a 2.4 GHz Wi-Fi network in their home, and slept in a bed at night (i.e., not a reclining chair or similar). Participants' bed partners were eligible to participate if their ASA PS class was II or lower and they slept in the same bed as the pregnant participant. Exclusion criteria for both the participant and their bed partner included non-English speaking, reading, or writing and ASA PS class III or higher.

For the participant eligibility criteria of the controlled-setting study, please refer to the publication.[5]

Interventions

Eligible participants and their bed partners provided voluntary, written, informed consent to participate. Basic demographic data was collected from each participant and their participating bed partner. Each participant used a home-surveillance camera to record a video of themselves and their bed partner sleeping overnight for a total of four nights. See **Appendix A** for additional details regarding interventions in the real-world study.

For the interventions in the controlled-setting study, see the publication.[5]

Outcomes

The main outcomes for this study were the precision and recall of the model for detecting thirteen pre-defined body positions (see**Dataset Development** , below). The outcome of the real-world study pertaining to the present study was the sleeping position of the pregnant participant and bed partner. See **Appendix A** for additional details. For the outcomes in the controlled-setting study, see the publication.[5]

Sample Size

For the real-world study, we selected a target sample size of N=60 couples (see **Appendix A**). Recruitment for the real-world study is ongoing, and the present study uses a subset of this data. In the controlled-setting study, the target sample size was twenty.[5]

Statistical Methods

We assessed the normality of continuous variables via the Shapiro-Wilk test at a 0.05 significance level and indicated deviations from normality in our results. Statistical analyses were conducted using the R Statistical Software package (Version 4.2.2).[6]

Dataset Development

All video recordings were reviewed manually by two trained reviewers and frames (images) were extracted at specific time points. The reviewers manually and independently annotated the extracted frames using an open-source annotation tool, LabelImg (see **Appendix A**).[7]Following our previous work,[5]body positions were classified into twelve position classes, including: left recovery, left lateral, left tilt, supine, supine thorax with left pelvic tilt, supine thorax with right pelvic tilt, supine pelvis with left thorax tilt, supine pelvis with right thorax tilt, right tilt, right lateral, right recovery, and sitting up at the edge of the bed (see **Figure A.1** in **Appendix A**). In addition, we added a prone position class for a total of thirteen body position classes.

Since pillows can impact sleeping positions, our model (SLeeP AIDePt-2) was designed to detect pillows. This involved identifying pillows – such as head pillows, body pillows, pregnancy pillows, and wedge pillows – in various locations on the bed. As a result, we thoroughly reviewed the controlled-setting dataset and annotated the presence of pillows in those frames as well (see **Appendix A**).

Model Development and Evaluation

We framed the detection of sleeping positions and pillows as a multi-participant, multi-class, classification problem where state-of-the-art deep Convolutional Neural Networks were used. In this regard, we used YOLOv5s (You Only Look Once, version 5s) with 7.5 million parameters,[8,9]including pre-trained weights from the COCO (Common Objects in Context) dataset and fine-tuned it on our annotated dataset to make predictions of classes. The experiment was implemented using the PyTorch framework. The model was trained using stochastic gradient descent with momentum, an initial learning rate of 0.01, a batch size of 32, and early stopping criteria.[10,11]The training was conducted under the Linux virtual machine running on Google Colab with an NVIDIA A100 (SXM4) 40GB GPU.

Our dataset was split into five non-overlapping folds, and we then trained, validated, and tested SLeeP AIDePt-2 by five-fold cross validation. In each loop of the cross-validation, three folds served as the training set, one fold was the validation set, and one fold was the testing set. Five loops were completed so each fold was given a turn to be the validation set and testing set. The weights that achieved the highest mean average precision (mAP) on the validation set were saved for each loop. Finally, performance evaluation was completed using the best weights on the testing set for each loop, and the following standard performance measures were calculated for each class: precision (analogous to "positive predictive value" in the clinical realm), recall (analogous to "sensitivity"), and average precision (AP), which is the area under the precision-recall curve (similar to how AUROC is the area under the receiver-operator curve), at standard intersection over union (IoU) values.

See **Appendix A** for more details regarding our rationale for using YOLOv5s and disabling flip augmentation; use of k-fold cross validation, early stopping criteria, a weighted random sampler; and definitions of standard performance measures, allocation of the study participants to the five folds, and the training, validation, and testing set assignments for each loop.

Patient and Public Involvement

Patients and public were not involved in the design of this study; however, this rationale for this study was informed by the experiences of pregnancy loss parents, who have brought maternal sleeping position to the attention of researchers and clinicians internationally.[12]

Results

From July 2022 through April 2023, 51 people expressed interest in participating in the real-world study. Of these, 34 (67%) were not assessed for eligibility (four decided against participating after learning more about the study, 25 did not respond after their initial expression of interest, and five gave birth prior to screening) – we did not collect any data from these. Of the seventeen (33%) participants that were screened, all met the eligibility criteria and gave written informed consent. All seventeen had bed partners, but only fifteen bed partners gave informed consent and participated. Two participants (and their bed partners) installed the camera incorrectly, so their data was excluded from the model building. As such, fifteen participants (and thirteen bed partners) successfully completed the study.

Demographic Characteristics

Demographic characteristics of the pregnant participants and their bed partners (if applicable) are shown in **Table 1** . Ethnic backgrounds from the real-world dataset included representation from Northern European, Latino, Greek, South Asian, East Asian, Armenian/Turkish, Italian, and Hungarian ancestries. We did not collect ethnicity, gravida, or parity data from the participants in the controlled-setting dataset.

Dataset

In the real-world study, we collected 29,253 minutes (487.6 hours) of infrared video from which we extracted and annotated 6,960 unique frames. Of these, 6,514 were multi-participant frames, and 446 were single-participant frames. See **Appendix B** for a qualitative description of the real-world dataset. These data from our real-world dataset were combined with our controlled-setting dataset (5,970 frames) for a total of 12,930 annotated frames, which contained 47,001 annotations, and comprised our dataset for building SLeeP AIDePt-2. See**Table 2** for class-wise information about the datasets.

A bar chart of the frequency of occurrences of each position class in the real-world dataset and controlled-setting dataset are shown in**Figure 1** . The sleeping positions have been rearranged on the x-axis of the bar chart to demonstrate the progression of the positions starting from left recovery and proceeding leftward, rolling across the back (supine), until right recovery and, finally, prone and sitting.

Models

In **Figure 2** , class-wise results (averaged across all five loops) are shown using a bar chart and heat map. The bar chart in**Figure 2A** shows the four performance metrics from the testing phase averaged across the five models' (one model per loop of the cross-validation) testing sets and across all classes. The error bars on the bar chart represent one standard deviation of the respective value across all measures, reflecting the variability across models (n=5) and classes (n=14). The heatmap in **Figure 2B** shows the four performance parameters (columns) from the testing phase averaged across the five models' test sets for each of the predicted classes (rows).

On a per-class basis and averaged across the five models, the sitting class had the highest AP@0.50 (0.92). The left lateral, right lateral, and supine classes also had high values of AP@0.50 (0.82 to 0.89), whereas recovery, prone, and twisted/hybrid positions generally had intermediate values of AP@0.50 (0.62 to 0.72). The non-hybrid tilted positions (left tilt and right tilt) had the lowest values of AP@0.50 (<0.50). As for the pillow class, the AP@0.50 was intermediate-to-high (0.80).

See **Table B.1** and **Table B.2** in **Appendix B** for a loop-wise summary of the training, validation, and performance testing of the cross-validation of SLeeP AIDePt-2.

A running example using one of our trained SLeeP AIDePt-2 models to localise and classify the sleeping position of a study participant and their bed partner in eight different extracted frames is displayed in**Figure**

3 .

Harms

There were no known or identified harms related to this study.

Discussion

Main Findings

We collected a heterogeneous video dataset containing instances of single participants and multiple participants, sleeping in or simulating twelve unique sleeping positions along with a sitting position, differing in sleeping environments and pregnancy status (third trimester, non-pregnant), with varying usage of pillows (head, body, pregnancy, wedge) and bed sheets (none, thin, thick; and pattern-free, patterned, striped, textured), both in infrared scale and colour scale video. We trained, validated, and tested a model (SLeeP AIDePt-2) on this dataset to detect the body positions and pillow use of a pregnant person and their bed partner (if any) appearing in an overnight video recording of sleep. The model best detects (high AP@0.50) pillows and the sitting, left lateral, right lateral, and supine positions and has a relatively low false negative rate (high recall) for these detections. The model detects less frequently occurring positions (prone, left/right recovery, supine thorax with left/right pelvic tilt, supine pelvis with left/right thorax tilt) less accurately and is particularly challenged by left/right tilt on which its performance is poorest.

Strengths and Limitations

This study describes the transition of a vision-based sleeping position detection model in preparation for real-world use in pregnancy research. The SLeeP AIDePt-2 model has many strengths. Notably, its real-world deployment does not require specialised equipment and takes into account low-lighting, bedsheets, and entry/exit events. It has learned factors unique to pregnancy anatomy and physiology such as determination of the pelvis position (supine, tilt, lateral, recovery) and direction (left, right). It is trained to detect multiple participants and other objects such as pillows in bed simultaneously, enabling it to not only account for more natural occurrences and frequencies of sleeping positions but also more natural sleeping contexts, behaviours, objects, and environments. Prone sleeping in late pregnancy has never been reported in the literature as naturally occurring and, as such, we believe it is exceedingly rare. However, we observed in our real-world dataset that prone sleeping is common in non-pregnant adults. SLeeP AIDePt-2 is trained to detect this. Compared to our previous work,[5]SLeeP AIDePt-2 is built on an expanded dataset containing a total of 52 participants (39 pregnant, 13 bed partners), including fifteen additional sleeping environments with no restrictions on bedsheets (thickness or patterns) and pillows.

This study has some limitations. While the real-world dataset contains many possible naturally occurring sleeping positions, SLeeP AIDePt-2 does not account for naturally occurring sleeping positions other than the twelve that we predefined. The overall sample size on which SLeeP AIDePt-2 is trained, particularly the real-world dataset, is small and could benefit from further increases in the number of participants, bed partners, and sleeping environments. The performance of SLeeP AIDePt-2 is sensitive to camera placement. Finally, we did not train SLeeP AIDePt-2 to detect household pets, which could impact sleeping position.[13,14]See **Appendix C** for further details.

Interpretation

Despite the current model (SLeeP AIDePt-2) being tested on a more challenging test set than our previous model (SLeeP AIDePt-1) (see**Appendix C**), it significantly outperformed SLeeP AIDePt-1 for the left lateral, supine, and right lateral positions with AP@0.50's of 0.89 (vs. 0.72), 0.82 (vs. 0.68), and 0.84

(vs. 0.64), respectively. In contrast, for almost all other positions (except sitting and supine pelvis with right thorax tilt), SLeeP AIDePt-2 performed slightly worse than SLeeP AIDePt-1. The explanation for this difference lies in the underlying datasets. The frequencies of occurrence of the sleeping positions in the controlled-setting dataset on which SLeeP AIDePt-1 was built were approximately equal (**Figure 1**); however, this was not so for the sleeping positions in the real-world dataset because people do not spend equal time sleeping in every possible position but, instead, shift between their two or three most comfortable positions. As such, despite our efforts to mitigate class imbalance during training, SLeeP AIDePt-2 is biassed to best detect the most frequent naturally occurring sleeping positions in pregnancy and in non-pregnant adults, which are left lateral, supine, and right lateral.

Our results are comparable to those from other vision-based sleeping position models in non-pregnant adults.[15–20]Both Li et al,[19]and Mohammadi et al,[17]used video recordings from a home-surveillance camera, leveraged controlled-setting data, accounted for bed sheets, and used a similar training methodology to us. Li et al, also combined real-world data with their controlled-setting data, as we did, to build their model.[19]Unfortunately, Li et al, did not present their model's performance results in a format that can be compared to ours (see **Appendix C**); however, comparing our model to Mohammadi et al,'s model, which was trained and validated on approximately 10,000 frames, we calculated their average recall (sensitivity) in the presence of bed sheets to be 0.64 (standard deviation 0.10), which is similar to ours (0.66, standard deviation 0.20). Beyond this, direct comparison of SLeeP AIDePt-2 to other vision-based models and position sensors is challenging because, as far as we are aware, no other measurement tool accounts for the positions of the pelvis and thorax simultaneously.

Conclusion

A computer vision algorithm for prospective measurement of sleeping position using video under real-world conditions across the third trimester in the home setting was built and may hold potential to solve yet unanswered research and clinical questions regarding the relationship between sleeping position and adverse pregnancy outcomes.

While it is currently unknown whether the association between maternal supine sleeping position and adverse pregnancy outcomes is causal, it is certain that there currently exists much controversy regarding the association itself.[4,21–25]The primary source of scepticism is the data underlying this association, which is based on subjective, self-reported recollection of sleeping position and does not account for intra-and inter-night variability in sleeping position.[1,2,4]As such, the current study makes a significant stride toward enabling the study of sleeping positions in the third trimester of pregnancy using more objective methods that were developed with the unique anatomy and physiology of pregnancy in mind and can be deployed using ubiquitous equipment without the need for the researcher to visit the participant's home. Using a model like SLeeP AIDePt-2, researchers could automate measurement of the time spent in each sleeping position across the third trimester, linking it to pregnancy outcomes. As is standard practice in machine learning and software applications, SLeeP AIDePt-2 should be continually audited, iterated, and updated by adding new data to the dataset and re-training, re-validating, and re-testing the model.[26]

For future research, we propose, "The DOSAGE Study: Dose Of Supine sleep Affects fetal Growth? an Exposure-response Study." The DOSAGE Study will be an international, prospective, cohort study aiming at gathering objective evidence that will either lend support to, or detract support from, a causal pathway between supine sleeping position after 28 weeks' gestation, foetal growth, and late stillbirth, and to quantify the safe "dose" of nightly supine sleeping time, if any. See**Appendix C** for further details.

Acknowledgments

None.

Disclosure of Interests

Allan Kember is the majority shareholder and the volunteer (unpaid) Chief Executive Officer and President of a startup company, Shiphrah Biomedical Inc. (SBI), which was one of the study funders. Allan Kember received financial payment from Mitacs through an entrepreneurship internship as part of his graduate studies at the University of Toronto (UofT). Min-En Hsieh received financial payment from Mitacs and UofT to complete an international exchange internship with SBI as part of his graduate studies at National Cheng Kung University. The other authors report no conflict of interest.

Contribution to Authorship

Allan Kember : Conceptualization, Methodology, Software, Validation, Formal analysis, Data Curation, Writing - Original Draft, Visualization, Project administration, Funding acquisition.**Min-En Hsieh** : Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Visualization. **Hafsa Zia** : Formal analysis, Investigation, Data Curation, Writing - Review & Editing. **Praniya Elangainesan** : Investigation, Data Curation, Writing - Review & Editing. **Ramak Adijeh** : Investigation, Data Curation, Writing - Review & Editing.**Ivan Li** : Methodology, Software, Formal analysis, Data Curation, Writing - Review & Editing. **Leah Ritchie** : Investigation, Data Curation, Writing - Review & Editing, Project administration.**Sina Akbarian** : Conceptualization, Methodology, Writing - Review & Editing. **Babak Taati** : Conceptualization, Methodology, Writing - Review & Editing, Supervision. **Sebastian Hobson** : Conceptualization, Methodology, Writing - Review & Editing, Supervision. **Elham Dolatabadi** : Conceptualization, Methodology, Resources, Writing - Review & Editing, Supervision, Funding acquisition.

Details of Ethics Approval

The data and procedures underlying this study came from two studies approved by the University of Toronto Health Sciences Research Ethics Board (Protocol No. 40985 approved 18JUN2021; Protocol No. 41612 approved 09MAY2022).

Funding

This study was funded by a Mitacs Entrepreneur-Accelerate Program grant (No. IT 26655). This program funds student and postdoctoral entrepreneurs to further develop the research or technology at the core of their start-up business by way of internships in collaboration with a university, professor and approved incubator. In this study, the total study funding was $15,000 CAD and was provided through Mitacs to the university (University of Toronto) for the professor (Elham Dolatabadi) to administer for the student/intern/entrepreneur (Allan Kember) and the study expenses related to the company's (SBI) technology under the supervision of the incubator (Health Innovation Hub). Mitacs provided 50% of the study funding. Mitacs' contribution to the grant was matched by SBI, which provided the remaining 50%. Mitacs had no role in the study design; collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication data. However, SBI with the oversight of the University of Toronto had a role in all these aspects via Allan Kember and Elham Dolatabadi.

References

1. Cronin RS, Li M, Thompson JMD, Gordon A, Raynes-Greenow CH, Heazell AEP, et al. An Individual Participant Data Meta-analysis of Maternal Going-to-Sleep Position, Interactions with Fetal Vulnerability, and the Risk of Late Stillbirth. EClinicalMedicine. 2019 Apr 1;10:49–57.

2. Anderson NH, Gordon A, Li M, Cronin RS, Thompson JMD, Raynes-Greenow CH, et al. Association of Supine Going-to-Sleep Position in Late Pregnancy With Reduced Birth Weight. JAMA Netw Open. 2019 Oct 2;2(10):e1912614.

3. Overview | Antenatal care | Guidance | NICE [Internet]. NICE; [cited 2022 Sep 2]. Available from: https://www.nice.org.uk/guidance/ng201

4. National Institute for Health and Care Excellence, National Guideline Alliance, Royal College of Obstetricians and Gynaecologists. Antenatal Care: [W] Maternal sleep position during pregnancy - NICE guideline NG201 - Evidence reviews underpinning recommendations 1.3.24 to 1.3.25 [Internet]. 2021 [cited 2022 May 24]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK573947/pdf/Bookshelf_NBK573947.pdf

5. Kember AJ, Selvarajan R, Park E, Huang H, Zia H, Rahman F, et al. Vision-based detection and quantification of maternal sleeping position in the third trimester of pregnancy in the home setting – building the dataset and model. PLOS Digit Health. 2023;In press.

6. The R Foundation for Statistical Computing. R Statistical Software. 2023.

7. Tzutalin. LabelImg Free Software [Internet]. 2015 [cited 2021 Sep 1]. Available from: https://github.com/heartexlabs/labelImg

8. Home · ultralytics/yolov5 Wiki [Internet]. GitHub. [cited 2022 Jun 7]. Available from: https://github.com/ultralytics/yolov5

9. YOLOv5 Documentation [Internet]. [cited 2022 Jun 7]. Available from: https://docs.ultralytics.com/

10. NVIDIA Tesla P100: The Most Advanced Data Center Accelerator [Internet]. NVIDIA. [cited 2022 Jun 7]. Available from: https://www.nvidia.com/en-us/data-center/tesla-p100/

11. Google Colaboratory [Internet]. [cited 2022 Jun 7]. Available from: https://research.google.com/colaboratory/faq.html

12. Warland J, O'Brien LM, Heazell AEP, Mitchell EA, STARS Consortium. An international internet survey of the experiences of 1,714 mothers with a late stillbirth: the STARS cohort study. BMC Pregnancy Childbirth. 2015 Aug 15;15:172.

13. Medlin K, Wisnieski L. The association of pet ownership and sleep quality and sleep disorders in United States adults. Hum-Anim Interact [Internet]. 2023 Mar 17 [cited 2023 Apr 3];2023. Available from: https://www.cabidigitallibrary.org/doi/10.1079/hai.2023.0005

14. Hoffman CL, Browne M, Smith BP. Human-Animal Co-Sleeping: An Actigraphy-Based Assessment of Dogs' Impacts on Women's Nighttime Movements. Anim Open Access J MDPI. 2020 Feb 11;10(2):278.

15. Grimm T, Martinez M, Benz A, Stiefelhagen R. Sleep position classification from a depth camera using Bed Aligned Maps. In: 2016 23rd International Conference on Pattern Recognition (ICPR). 2016. p. 319–24.

16. Liu S, Yin Y, Ostadabbas S. In-Bed Pose Estimation: Deep Learning with Shallow Dataset. IEEE J Transl Eng Health Med. 2019;7:1–12.

17. Mohammadi SM, Alnowami M, Khan S, Dijk DJ, Hilton A, Wells K. Sleep Posture Classification using a Convolutional Neural Network. Annu Int Conf IEEE Eng Med Biol Soc IEEE Eng Med Biol Soc Annu Int Conf. 2018 Jul;2018:1–4.

18. Wang YK, Chen HY, Chen JR. Unobtrusive Sleep Monitoring Using Movement Activity by Video Analysis. Electronics. 2019 Jul;8(7):812.

19. Li YY, Wang SJ, Hung YP. A Vision-Based System for In-Sleep Upper-Body and Head Pose Classification. Sensors. 2022 Mar 4;22(5):2014.

20. Akbarian S, Delfi G, Zhu K, Yadollahi A, Taati B. Automated Non-Contact Detection of Head and Body Positions During Sleep. IEEE Access. 2019;7:72826–34.

21. Silver RM, Hunter S, Reddy UM, Facco F, Gibbins KJ, Grobman WA, et al. Prospective Evaluation of Maternal Sleep Position Through 30 Weeks of Gestation and Adverse Pregnancy Outcomes. Obstet Gynecol. 2019 Oct;134(4):667–76.

22. McCowan LME, Cronin RS, Gordon A, O'Brien L, Heazell AEP, CRIBBS Collaboration. Prospective Evaluation of Maternal Sleep Position Through 30 Weeks of Gestation and Adverse Pregnancy Outcomes. Obstet Gynecol. 2020 Jan;135(1):218.

23. Silver RM, Reddy UM, Gibbins KJ. Response to Letter. Obstet Gynecol. 2020 Jan;135(1):218–9.

24. Fox NS, Oster EF. The Advice We Give to Pregnant Women: Sleep On It. Obstet Gynecol. 2019 Oct;134(4):665.

25. McClure E, Goldenberg R. More evidence needed for association between sleep practices and stillbirth. BJOG Int J Obstet Gynaecol. 2018;125(2):263–263.

26. Granlund T, Stirbu V, Mikkonen T. Towards Regulatory-Compliant MLOps: Oravizio's Journey from a Machine Learning Experiment to a Deployed Certified Medical Product. SN Comput Sci. 2021 Jun 19;2(5):342.

Table and Figure Caption List

Table 1. Demographic Characteristics of Participants in the Combined Dataset (Controlled-Setting and Real-world), Controlled-Setting Position Dataset, and Real-world Dataset

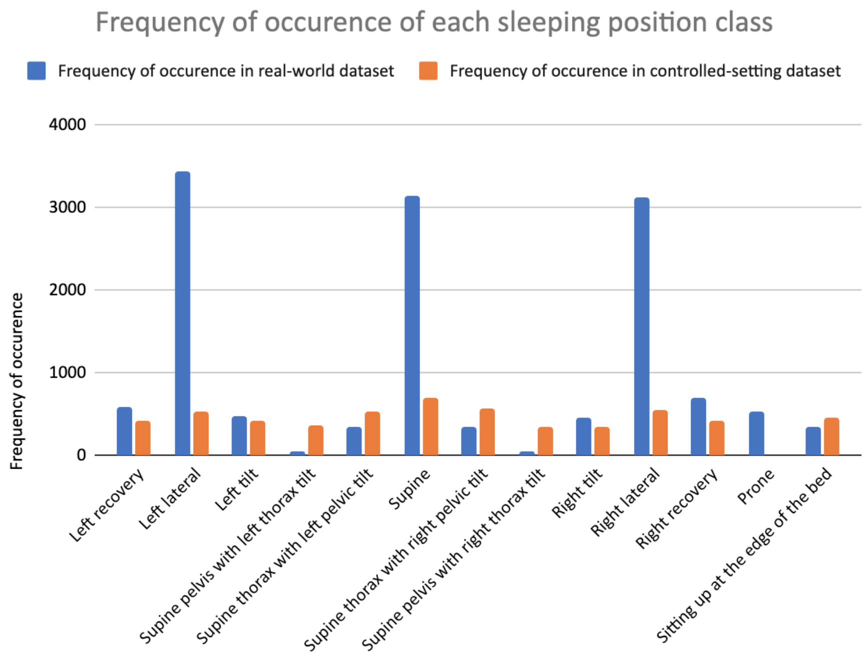
Table 2. Total Number of Annotations Containing Each Class in the Real-world Dataset, Controlled-Setting Dataset, and Combined (Controlled-Setting and Real-world) Dataset on Which SLeeP AIDePt-2 was Trained

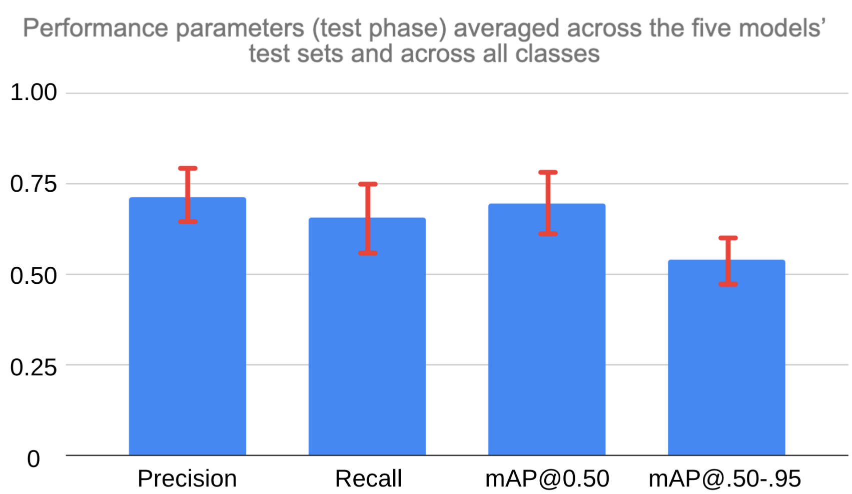
Figure 1. Bar Chart of the Frequency of Occurrences of Each Sleeping Position Class and the Sitting Class in the Real-world Dataset and Controlled-Setting Dataset. *Legend* : Real-world dataset shown in blue, and controlled-setting dataset shown in orange.

Figure 2A. Bar Chart of SLeeP AIDePt-2 Performance Metrics From the Testing Phase Averaged Across the Five Models' Test Sets and Across All Classes. *Legend* : mAP@0.50 indicates the mean average precision at an intersection of union of 0.50. mAP@.50-.95 indicates the mean average precision at intersections of unions between 0.50 and 0.95. The error bars represent one standard deviation of the respective value across all measures, which reflects the variability both across models and classes. The y-axis does not have units because precision, recall, mAP@0.50, and mAP@.50-.95 are dimensionless values.

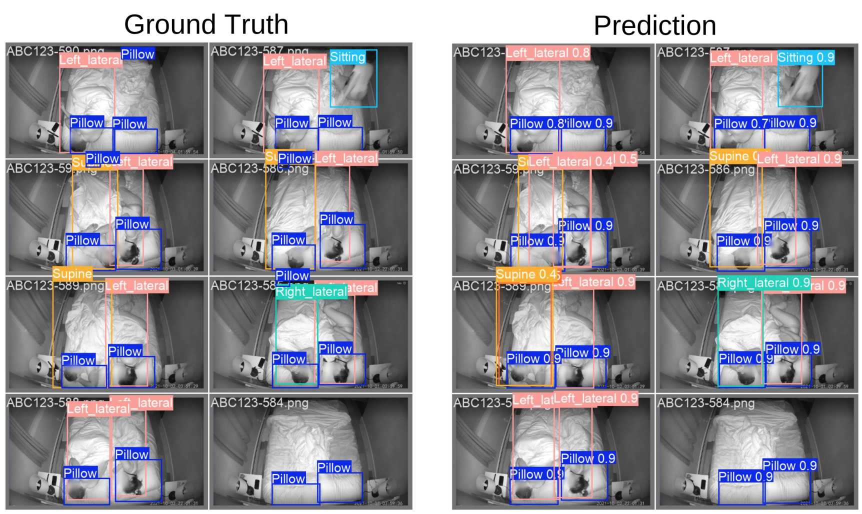
Figure 2B. Heatmap of SLeeP AIDePt-2 Performance Metrics (Columns) From the Testing Phase Averaged Across the Five Models' Test Sets For Each of the Predicted Classes (Rows). *Legend* : AP@0.50 indicates the average precision at an intersection of union of 0.50. AP@.50-.95 indicates the average precision at intersections of unions between 0.50 and 0.95. The value of the respective performance metric is mapped to a colour spectrum from red to yellow to green where values of 0.50 or less are represented by red at the lower end of the spectrum, values around 0.75 are shades around yellow (oranger if lower than 0.75; greener if higher than 0.75), and values of 0.90 or more are represented by green at the higher end of the spectrum. The "all body position classes average" is provided as the averaged value of the respective performance metric across the five models' test sets and the 13 body position classes. For the "all body position classes average" row, the value in the AP@0.50 column is the mean AP@0.50, and the value in the AP@.50-.95 column is the mean AP@.50-.95 since these values represent averages across multiple classes.

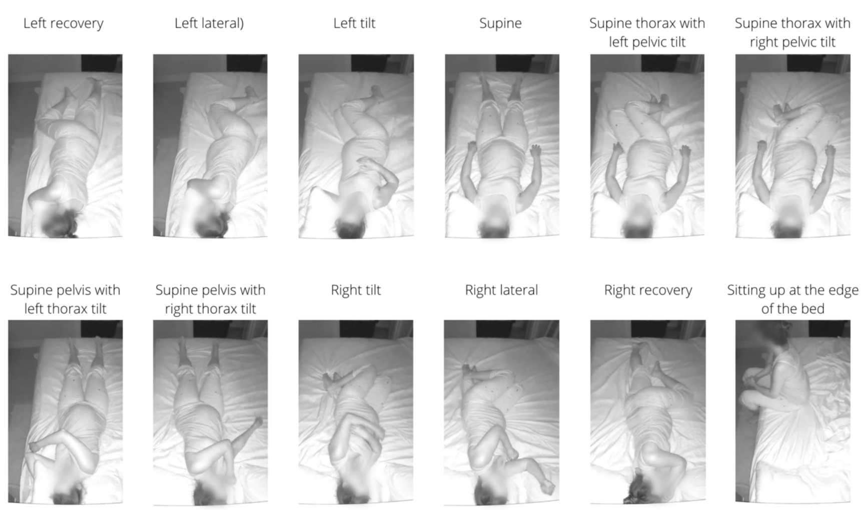
Figure 3. Example Output of SLeeP AIDePt-2 Localising and Classifying the Sleeping Positions of a Study Participant and Their Bed Partner as Well as Their Pillows in Eight Different Extracted Frames. *Legend* : The participant's and bed partner's body and pillow annotations are shown by the coloured boxes ("Ground Truth", left). The model's localization and prediction of the sleeping positions and pillows, along with its confidence score (between 0 and 1) at the top of each bounding box, are shown ("Prediction", right).

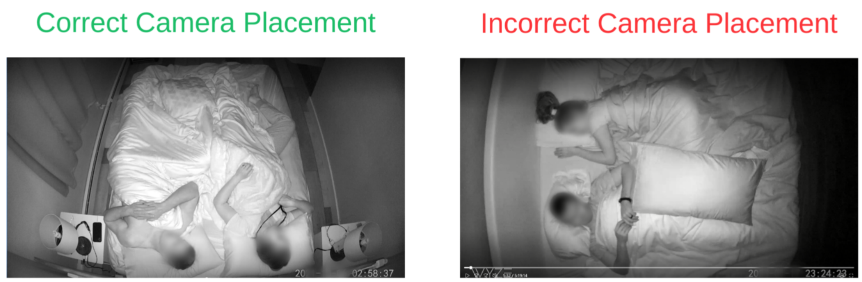




Performance parameter (test phase)	Precision	Recall	AP@0.50	AP@.50-.95
All classes average	0.71	0.66	0.70	0.54
Left recovery	0.66	0.62	0.66	0.51
Left lateral	0.80	0.85	0.89	0.61
Left tilt	0.53	0.40	0.45	0.34
Supine	0.70	0.81	0.82	0.61
Supine thorax with left pelvic tilt	0.71	0.58	0.62	0.55
Supine thorax with right pelvic tilt	0.65	0.59	0.64	0.55
Supine pelvis with left thorax tilt	0.75	0.66	0.69	0.59
Supine pelvis with right thorax tilt	0.69	0.64	0.69	0.61
Right tilt	0.54	0.33	0.41	0.32
Right lateral	0.77	0.82	0.84	0.57
Right recovery	0.63	0.69	0.72	0.55
Prone	0.61	0.69	0.68	0.49
Sitting	0.95	0.87	0.92	0.66
All body position classes average	0.69	0.66	0.69	0.54
Pillow	0.89	0.70	0.80	0.61







Hosted file

`Tables Document.docx` available at https://authorea.com/users/655602/articles/661329-transitioning-sleeping-position-detection-in-late-pregnancy-using-computer-vision-from-controlled-to-real-world-settings-an-observational-study