Deep Transfer Learning for Visually Induced Motion Sickness Detection Using Symmetric Projection Attractor Reconstruction of the Electrocardiogram

Emmanuel Molefi, Ramaswamy Palaniappan

University of Kent, Canterbury, UK

Abstract

Despite the ubiquity of motion sickness – a long recognized syndrome from ancient days of sea travel – it still remains a persistent problem. In fact, around one in three individuals can be severely susceptible to this malady. The electrocardiogram (ECG) is an essential tool that has long been used to examine the physiological expression of motion sickness; commonly by performing analysis of ECG-derived heart rate variability (HRV). Here, we obtained the symmetric projection attractor reconstruction (SPAR) transforms of ECG signals recorded from healthy participants at rest and during nausea, for a binary image classification task using a set of pretrained deep neural networks with transfer learning. Our observations provide new insights into how physiologic characteristics captured via ECG-derived attractor images may be important for the detection of ECG signals that show differential response to motion-induced nausea.

1. Introduction

Nausea is the hallmark of motion sickness – a polysymptomatic condition that arises from conflicts caused by ambiguous sensory information from brain systems governing motion perception (i.e., visual, vestibular, and proprioceptive) [1, 2]. These sensory conflicts or neural mismatches [3] occur, for instance, because the brain receives information about motion that our eyes cannot perceive (e.g., a bumpy car ride); or that visual input from the eyes communicates movement that our bodies cannot detect (e.g., virtual reality applications).

Besides nausea and vomiting, motion sickness can also initiate symptoms, such as sweating, dizziness, and drowsiness; including perturbations in autonomic nervous system (ANS) function. In fact, because motion sickness can disrupt ANS state, i.e., reducing parasympathetic neural response while promoting sympathetic activation [4], the electrocardiogram (ECG) is an oft-used tool to study the response of individuals presented with motion sickness-inducing stimuli. Previous studies have commonly utilized the ECG as a conduit for heart rate variability (HRV) signal derivation – with subsequent evaluation of ANS functional state. This indirect approach of using the ECG is susceptible to important ECG waveform profile information being underutilized [5].

To glean insights into ECG morphological changes in response to motion-induced nausea, we leverage a recently developed mathematical tool – symmetric projection attractor reconstruction (SPAR) [6, 7] – that transforms any approximately periodic signal into a two-dimensional (2D) image ("attractor"); providing visualization and quantification of, e.g., ECG morphology and variability. Because attractor analysis and interpretation is mostly a visual task, we thus investigate the potential of convolutional neural network (CNN)-based deep learning architectures to determine or detect whether ECG records of participants during states of "baseline" (i.e., rest) and "nausea" characterize with distinct ECG profiles (i.e., differences in ECG morphology and variability). Specifically, here we utilize transfer learning, a widely used technique in deep learning applications – for example, in computer vision – that enables performing classification on a new task by customizing or fine-tuning pretrained deep neural networks. Prior research on biomedical image classification has applied this learning technique for, e.g., heart disease [8], breast cancer [9], and COVID-19 [10] detection. In addition, Aston *et al*. [11] demonstrated the potential for deep transfer learning application toward classification of ECG signals from wild-type and $Scn5a^{+/-}$ mutant mice via SPAR.

Herein, we sought to explore whether machine models can differentiate normal ECG-derived attractors from those showing signs of malaise; the significance of which could be to aid early detection of motion-induced nausea onset in real-time for timely therapeutic intervention.

2. Methods

2.1. Dataset

We consider ECG recordings from an experimental platform designed to induce motion sickness utilizing a nauseogenic visual stimulus. Extended details of the nau-

seogenic stimulation and dataset acquisition have been previously described in [12]. The dataset – comprising a combined 6 min duration at "baseline" and "nausea" states – were obtained from 12 participants (mean age 26.8 years; range 21-34 years; 10 female). All protocols were approved by the University of Kent Central Research Ethics Advisory Group (ref: CREAG015-12-2021), and conformed to the principles of the Declaration of Helsinki; all participants provided written informed consent.

2.2. SPAR Analysis

The ECG records were extracted into 18 epochs (each of 10 s duration) per state (i.e., "baseline" and "nausea" states) and then transformed to attractor images using SPAR. Prior to SPAR analysis, the ECG data were detrended, smoothed with a Savitzky-Golay filter, and range normalized to the interval $[0, 1]$. The SPAR method has been well-described by authors in [6], and further extended to higher dimensional embedding by Lyle and Aston [7]. Briefly, the original SPAR used $N = 3$ equally spaced points on the time series signal to perform phase space reconstruction based on Takens' delay coordinate embedding [13]. Thus, to reconstruct a three-dimensional (3D) phase space for an ECG time series $x(t)$, further delay coordinates are given by

$$y(t) = x(t - \tau), \qquad z(t) = x(t - 2\tau), \tag{1}$$

where time delay τ is defined as $1/N$ of the average cycle length (i.e., cardiac cycle duration – determined using the Pan-Tompkins algorithm [14]). Next, to remove baseline shift (e.g., which may result from respiration and movement), the 3D reconstructed attractor is projected onto a plane (v, w) that is orthogonal to the vector $[1, 1, 1]$ via

$$v = \frac{1}{\sqrt{6}}(x + y - 2z), \qquad w = \frac{1}{\sqrt{2}}(x - y). \tag{2}$$

Herein, to generate attractors for $N \geq 3$, we use the generalized SPAR method which computes an N-dimensional phase space embedding. In this general case, an ECG time series $x(t)$ with period T is embedded into $N \geq 3$ dimensions using the time delay $\tau = T/N$ by the coordinates $x_{N,j}(t) = x(t - j\tau),\ j = 0, ..., N - 1$; next, [7] define the $a_{N,k}(t)$ and $b_{N,k}(t)$ coordinates as

$$\begin{aligned} a_{N,k}(t) &= \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} \cos(2\pi jk/N) x_{N,j}(t), \\ b_{N,k}(t) &= -\frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} \sin(2\pi jk/N) x_{N,j}(t) \end{aligned} \tag{3}$$

for $k = 1, ..., \lfloor (N - 1)/2 \rfloor$. Of note, here we use $k = 1$ for all embedding projections. Figure 1 shows example attractors generated via this implementation.

2.3. Transfer Learning

Because of our small dataset (i.e., $n = 12$), we use transfer learning; and consider the pretrained neural networks DenseNet-201, ResNet-50, ResNet-101, Xception, Inception-v3, and Inception-ResNet-v2, trained on ImageNet [15]. We provide as input, attractor images from the two states ("baseline" and "nausea"); example attractors are shown in Figure 1 for both states. There are 18 attractor images for each state, respectively; thus, each participant has 36 images, resulting in a total of 432 images from all participants. Using augmented image datastores, input images were automatically resized to the image input size of the respective network for compatibility (e.g., 224×224 for DenseNet-201). Because the last few layers (i.e., network head) of these networks are configured for 1000 classes, we fine-tune these final layers for the new binary classification task ("baseline" and "nausea" classes); replacing the last layer with learnable weights with a new fully connected layer with an output size of 2 (setting the learning rate factor for weights and biases on this layer to 10; to learn faster). We replace the classification layer with a new one without class labels. During model training, the *trainNetwork* function in MATLAB R2023b (The MathWorks, Inc., Natick, MA, USA) automatically sets the output classes of the layer.

To assess model generalization performance to unseen participants, we performed a leave-one-participant-out cross-validation (LOPOCV), where we left out all images of a single participant from the training images. This helps avoid biased classification accuracy. We utilized Bayesian optimization with 100 iterations to tune the hyperparameters mini-batch size, learning rate, momentum, and L2 regularization using the SGDM solver. All experiments were performed in MATLAB with a single 80GB NVIDIA A100 GPU.

3. Results

We observed that the DenseNet-201 network performed better than other networks as evaluated by accuracy for attractors generated using $N = 3, 5,$ and 6 points (Table 1). When examining model performance using the receiver operating characteristic (ROC) curve and the area under a ROC curve (AUC), we found that ResNet-50 and ResNet-101 achieved AUC scores considered acceptable for 4-point attractors (AUC 0.7091 and AUC 0.7275, respectively; Figure 2a,b). The DenseNet-201 model demonstrated acceptable AUCs for attractors computed using embedding dimensions ($N = 3$; AUC 0.7325) and ($N = 6$; AUC 0.7098) (Figure 2c).

Attractor appearance evaluation via visual examination reveals that there are distinctive features during motion sickness-induced nausea (Figure 1; Nausea panels) com-

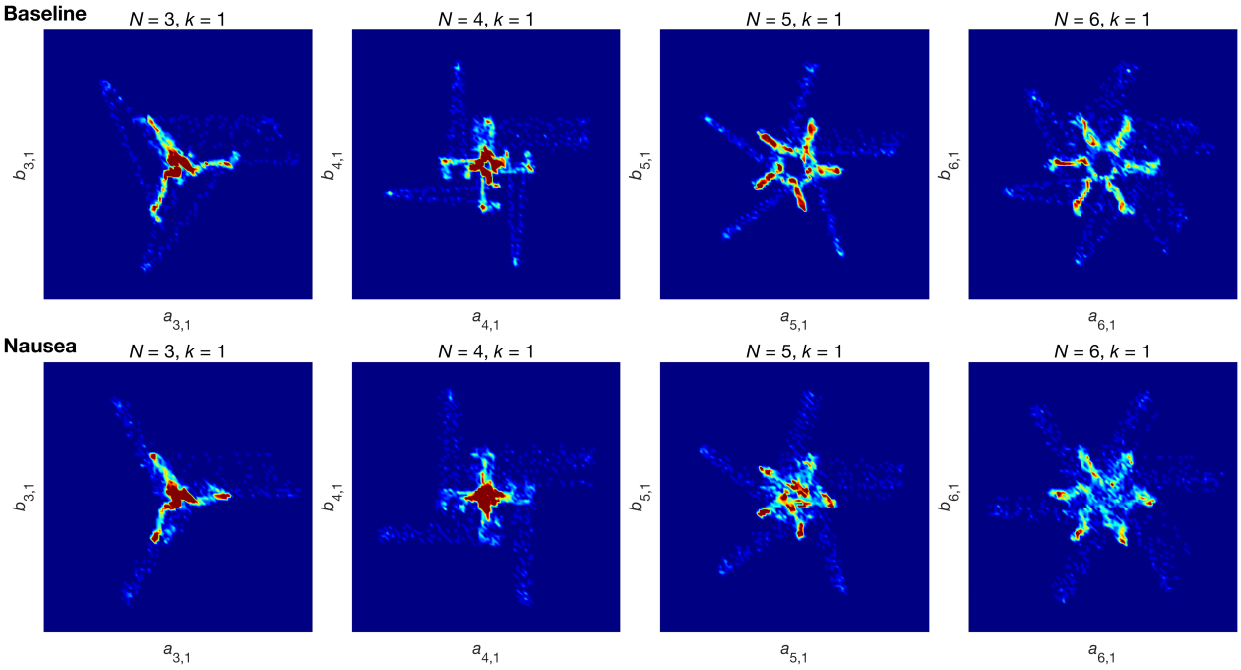

Figure 1. Example attractors for embedding dimensions $N = 3, ..., 6$, generated from electrocardiogram (ECG) signals of a participant at baseline (i.e., rest) and during nausea states.

pared to rest (Figure 1; Baseline panels). Particularly, the arms of the $(3, 1)$ attractor at baseline are wider than those of the nausea $(3, 1)$ attractor (Figure 1). Further, we can clearly see a much greater visual distinction between baseline and nausea $(5, 1)$ attractors, portraying how baseline attractors are less dense near the center than their nausea counterparts (Figure 1). Interestingly, these $(5, 1)$ attractor visual differences may explain the performance demonstrated by the DenseNet-201 network for $N = 5$ embedding, where it achieved 70.14% accuracy; whereas via majority vote, obtained 79.17% accuracy (Table 1).

Table 1. Pretrained neural networks results for binary classifications at four attractor embedding dimensions, presented as accuracy % (top), majority vote % (bottom).

Network	**3**	**4**	**5**	**6**
ResNet-50	67.13	69.21	68.52	68.98
	70.83	79.17	75.00	83.33
ResNet-101	69.21	70.37	65.51	67.59
	79.17	79.17	66.67	66.67
DenseNet-201	70.14	67.59	70.14	69.91
	75.00	70.83	79.17	79.17
Xception	65.05	66.44	62.50	65.28
	75.00	66.67	66.67	66.67
Inception-v3	67.82	67.13	62.04	66.20
	87.50	83.33	66.67	79.17
Inception-ResNet-v2	66.67	68.52	65.28	69.68
	70.83	83.33	66.67	70.83

4. Discussion

We provide insights into the potential of objectively identifying individuals during a nauseogenic experience using deep learning via a set of experiments utilizing pretrained neural networks – for classification of attractor images. Moreover, visual examination of these attractors (Figure 1) – which portray "hidden" characteristics in ECG morphology and variability and, by extension, HRV [16] – suggests differential response to motion-induced nausea.

While some pretrained neural networks achieved acceptable AUC values, e.g., DenseNet-201, we note, however, that these networks were not developed to specifically handle motion-induced malaise attractor images. Yet this suggests a way in which motion sickness could be detected or predicted using attractor-tailored deep learning algorithms, which in turn, may enable early detection of symptom onset; with implications for timely pharmacological or non-pharmacological (e.g., neuromodulation) therapeutic intervention. Furthermore, the attractor differences observed here suggest encapsulated ECG morphological characteristics that may serve as potential targets for motion sickness management. Interestingly, previous research has implicated ECG attractor images with drug treatment effects evaluation [17]. Altogether, our findings suggest that using SPAR may offer a new way of examining motion-induced nausea, and the opportunity to build algorithms that could effectively detect or predict it.

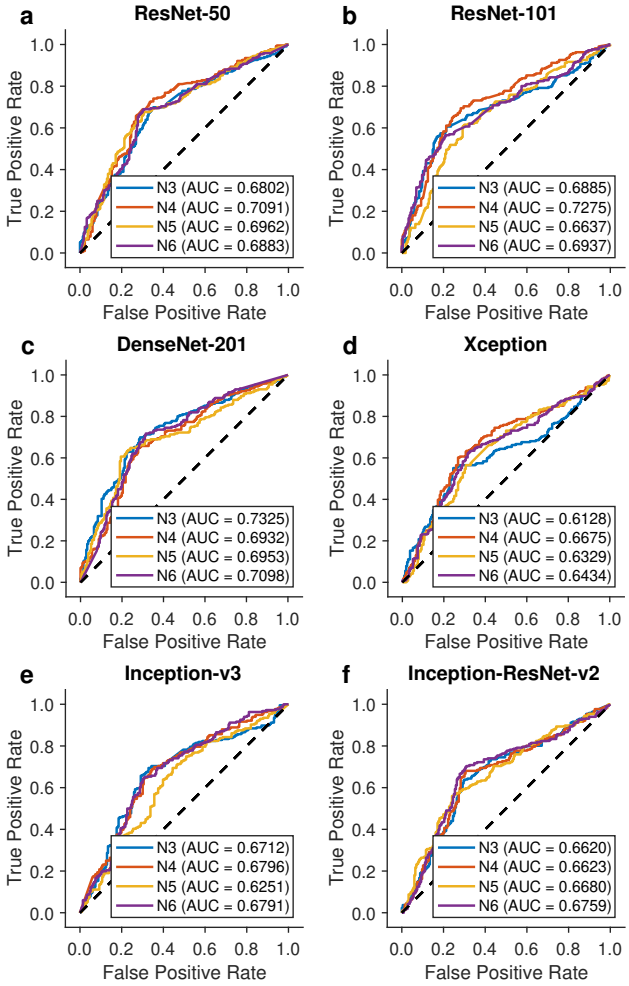

Figure 2. **(a-f)** ROC curves of all explored pretrained deep neural networks and embedding dimensions showing the true positive rate (TPR), or sensitivity, versus the false positive rate (FPR), or 1-specificity, for the "nausea" class.

5. Conclusion

ECG-derived attractors show promise for detection of differential response to motion-induced nausea using deep neural networks; this implies that manual features obtained via SPAR transformation may help quantify and assess the efficacy of new therapeutics for motion sickness management. Further research is needed to examine the wider role of SPAR in understanding and detecting motion sickness effects on autonomic function.

Acknowledgments

E.M. was supported by an Engineering and Physical Sciences Research Council (EPSRC) PhD studentship at the University of Kent School of Computing (EP/T518141/1).

References

[1] Reason J, Brand J. Motion Sickness. Academic Press, 1975.

[2] Oman CM. Motion sickness: a synthesis and evaluation of the sensory conflict theory. Can J Physiol Pharmacol 1990; 68(2):294–303.

[3] Reason JT. Motion sickness adaptation: a neural mismatch model. J R Soc Med 1978;71(11):819–829.

[4] Kim J, Napadow V, Kuo B, Barbieri R. A combined HRV-fMRI approach to assess cortical control of cardiovagal modulation by motion sickness. In Proc IEEE EMBS. IEEE, 2011; 2825–2828.

[5] Lyle JV, Charlton PH, Bonet-Luz E, Chaffey G, Christie M, Nandi M, Aston PJ. Beyond HRV: Analysis of ECG signals using attractor reconstruction. In Proc CinC. 2017; 1–4.

[6] Aston PJ, Christie MI, Huang YH, Nandi M. Beyond HRV: attractor reconstruction using the entire cardiovascular waveform data for novel feature extraction. Physiol Meas 2018;39(2):024001.

[7] Lyle J, Aston P. Symmetric projection attractor reconstruction: Embedding in higher dimensions. Chaos 2021;31(11).

[8] Tekin H, Kaya Y. A new approach for heart disease detection using Motif transform-based CWT's time-frequency images with DenseNet deep transfer learning methods. Biomed Tech Berl 2024;.

[9] Rani N, Gupta DK, Singh S. Multi-class classification of breast cancer abnormality using transfer learning. Multimed Tools Appl 2024;1–16.

[10] Hassan E, Shams MY, Hikal NA, Elmougy S. Detecting COVID-19 in chest CT images based on several pre-trained models. Multimed Tools Appl 2024;1–21.

[11] Aston PJ, Lyle JV, Bonet-Luz E, Huang CL, Zhang Y, Jeevaratnam K, Nandi M. Deep Learning Applied to Attractor Images Derived from ECG Signals for Detection of Genetic Mutation. In Proc CinC, volume 46. 2019; 1–4.

[12] Molefi E, McLoughlin I, Palaniappan R. On the potential of transauricular electrical stimulation to reduce visually induced motion sickness. Sci Rep 2023;13(1):3272.

[13] Takens F. Detecting strange attractors in turbulence. Dynamical Systems and Turbulence Warwick 1980 Lecture Notes in Mathematics 1981;898:366–381.

[14] Pan J, Tompkins WJ. A Real-Time QRS Detection Algorithm. IEEE Trans Biomed Eng 1985;BME-32(3):230–236.

[15] Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L. ImageNet: A large-scale hierarchical image database. In Proc IEEE CVPR. 2009; 248–255.

[16] Molefi E, McLoughlin I, Palaniappan R. Heart Rate Variability Responses to Visually Induced Motion Sickness. In Proc IEEE EMBS. IEEE, 2023; 1–4.

[17] Lyle JV, Nandi M, Aston PJ. Investigating the Response to Dofetilide with Symmetric Projection Attractor Reconstruction of the Electrocardiogram. In Proc CinC, volume 46. 2019; 1–4.

Address for correspondence:

Emmanuel Molefi
School of Computing, University of Kent
Canterbury, Kent, CT2 7FS, UK
em576@kent.ac.uk