Scaling Dynamic Mode Decomposition For Real Time Analysis Of Infant Movements

Navya Annapareddy, Lisa Letzkus, Santina Zanelli, Stephen Baek

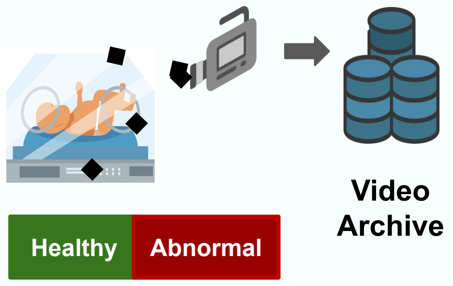
UVA Children's
UVA SCHOOL of DATA SCIENCE

Introduction

- Early detection of neuromotor disorders in clinical contexts rely on identifying subtle, transient spatiotemporal movement patterns often obscured in high-dimensional pose data.
- We explore the use of Dynamic Mode Decomposition (DMD) to classify infant motion risk. DMD separates time-evolving dynamics from spatial modes with distinct frequency and growth rates.
- Unlike PCA, which collapses spatial and temporal information into static orthogonal components, these modes serve as interpretable and physiologically meaningful representations of movement patterns.

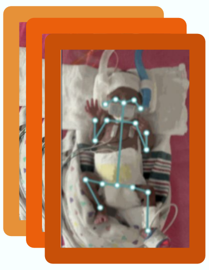
2D and 3D Pose Estimation

Video Movement Assessment



2D Pose Estimation Pipeline

joint	x	y
right elbow	…	…
…	…	…
right ankle	…	…



3D Pose and Depth Estimation

CNN Model	Backbone	Params	GFLOPs	Precision	Recall
SimpleBaseline	ResNet-50	32.42M	20.23	82.4	83.8
SimpleBaseline	ResNet-50	32.42M	20.23	90.1	91.6
DarkPose	HRNet-W48	60.65M	32.88	88.5	90.0
DarkPose + FiDIP	**HRNet-W48**	**60.65M**	**32.88**	**93.6**	**94.6**
Pose-MobileNet	**MobileNetV2**	**3.91M**	**0.46**	**76.5**	**67.2**

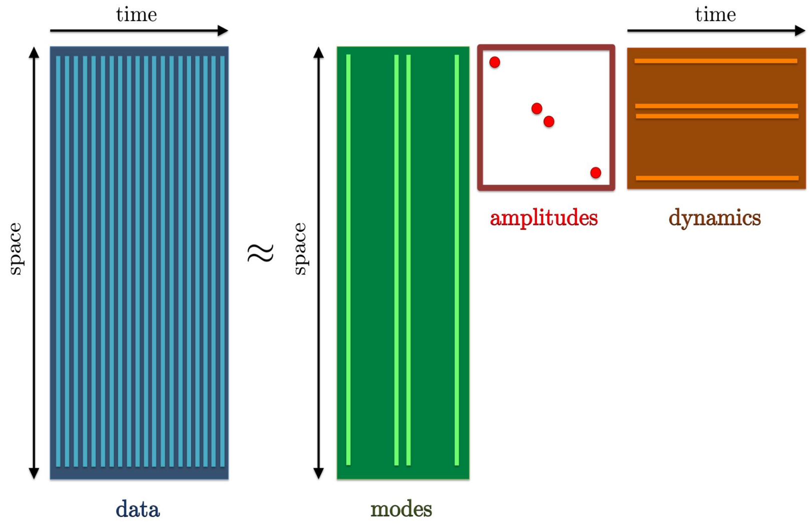
Dynamic Mode Decomposition

Real Time Anatomical Keypoints

$$\mathbf{Frame}_i = \begin{bmatrix} x_{1i} & y_{1i} \\ x_{2i} & y_{2i} \\ \vdots & \vdots \\ x_{17i} & y_{17i} \end{bmatrix}$$

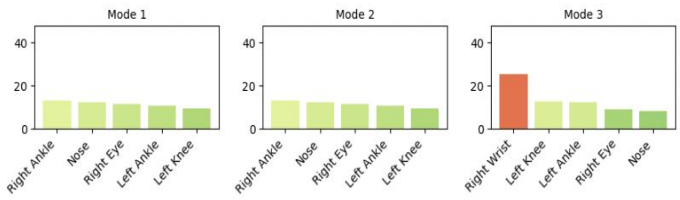
$$X = \begin{bmatrix} | & & | \\ x_1 & \cdots & x_m \\ | & & | \end{bmatrix} \qquad Y = \begin{bmatrix} | & & | \\ y_1 & \cdots & y_m \\ | & & | \end{bmatrix}$$

time

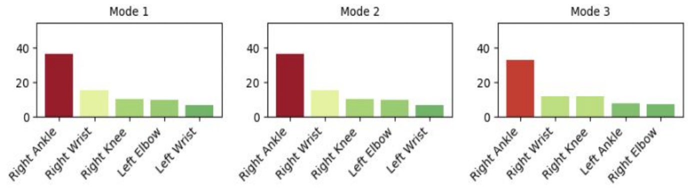


Results

Contributed Joint Variances (Healthy Infants)



Contributed Joint Variances (Abnormal Infants)



Classifier Model	Accuracy (%)			
	Healthy	Abnormal	Poor Repertoire	Overall
Autoencoder (AE)	93.2	80.1	75.4	86.3
Variational Autoencoder (VAE)	97.1	94.6	91.2	**95.0**
SVM (RBF Kernel)	92.5	85.3	79.8	88.9

Results utilized DarkPose estimator and GPU acceleration (A100) with an equally proportionate split across the 3 classes, with 220,000 training frames and 55,000 test frames.

Conclusions

- DMD reveals physically meaningful, temporally structured patterns for tasks involving dynamic classification, in a manner much more interpretable than traditional dimensionality reduction methods.
- The method's interpretability and alignment with real-time dynamics make it a promising tool for clinical screening, despite challenges like mode redundancy and sensitivity to noise.
- Future directions include extending to multi-class classification, refining mode selection via clustering.