LuKAN: A Kolmogorov-Arnold Network Framework for 3D Human Motion Prediction

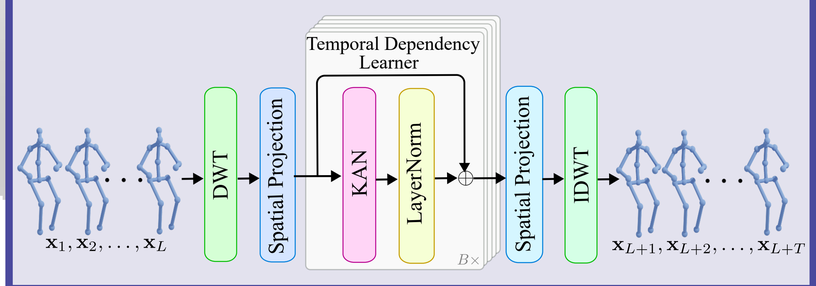
Md Zahidul Hasan, A. Ben Hamza, and Nizar Bouguila
Concordia University, Montreal, Canada

Problem

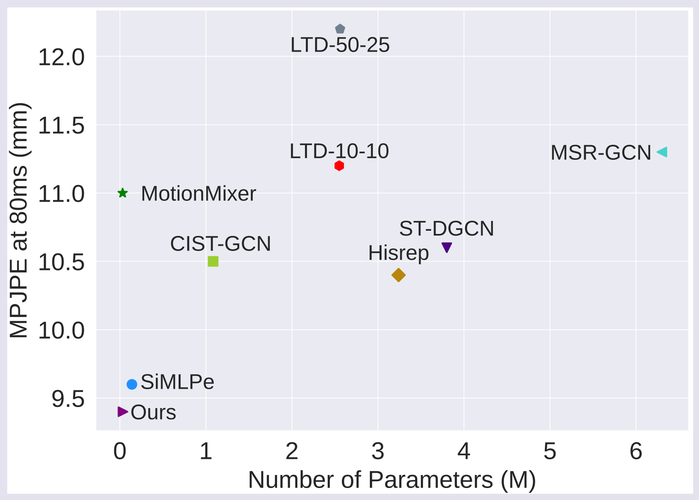
Let $\mathbf{X}_{1:L} = (\mathbf{x}_1, \ldots, \mathbf{x}_L)^\intercal \in \mathbb{R}^{L \times K}$ be a history motion sequence of L consecutive 3D human poses, where L is the look-back window, $K = 3J$ in the feature dimension, and J is the total number of body joints. At each time step t, each pose $\mathbf{x}_t \in \mathbb{R}^{1 \times K}$ is a flattened vector formed by concatenating the 3D coordinates of all joints in a single frame.

The objective is to construct a predictive model that estimates a motion sequence $\hat{\mathbf{X}}_{L+1:L+T} = (\hat{\mathbf{x}}_{L+1}, \ldots, \hat{\mathbf{x}}_{L+T}) \in \mathbb{R}^{T \times K}$ for the subsequent T timesteps. To this end, we design an efficient model based on Kolmogorov-Arnold networks.

Model Architecture



Model Efficiency Analysis



Method	MPJPE (mm)↓	Params (M)	FLOPs (M)
MSR-GCN	112.9	6.3	192.4
ST-DGCN	109.7	1.74	55.8
SiMLPe	109.4	0.14	8.76
CIST-GCN	110.3	1.048	49.7
LuKAN (Ours)	109.3	0.045	2.1

Time Complexity. Each spatial projection involves a matrix multiplication of complexity $\mathcal{O}(DJL)$, where J is the number of joints, D is the embedding dimension, and L is the length of the input sequence. DWT and its inverse are applied along the temporal dimension. As these are linear-time operations per sequence and per feature, their total complexity is $\mathcal{O}(JL)$. The core component of LuKAN is a B-layer KAN with Lucas polynomial activations, where B is the total number of blocks. Its time complexity is $\mathcal{O}(BDRL^2)$, where R is the degree of the Lucas polynomial. Hence, the time complexity of LuKAN is $\mathcal{O}(DJL + BDRL^2)$.

Memory Complexity. Our model maintains a lightweight parameter count. Each spatial projection require $\mathcal{O}(JD)$ parameters, while the B-layer KAN contributes $\mathcal{O}(BRL^2)$ parameters, giving a total parameter complexity of $\mathcal{O}(JD + BRL^2)$. During runtime, memory is also allocated for storing intermediate activations and for evaluating the polynomial basis, yielding a total runtime memory complexity of $\mathcal{O}(DL + JL)$. Overall, LuKAN achieves a compelling balance between expressive power and computational efficiency.

Method

Temporal Encoding. To encode temporal information of the human motion in the trajectory $\mathbf{x}^{(i)} = (x_1^{(i)}, \ldots, x_L^{(i)})^\intercal$, we employ DWT, which decomposes a signal into its approximate and detail components using wavelets, yielding $\mathbf{c}^{(i)} = \mathrm{DWT}(\mathbf{x}^{(i)})$, where $\mathbf{c}^{(i)} = (\mathbf{a}^{(i)}, \mathbf{d}^{(i)})^\intercal$ is an $(L_a + L_d)$-dimensional vector of wavelet coefficients (i.e., approximation and detail components).

Spatial Projection. The spatial projection maps the DWT-transformed history motion sequence into an embedding space of dimension D, capturing inter-joint dependencies. Its output is an $(L_a + L_d) \times D$ matrix given by $\mathbf{Z}_1 = \mathrm{DWT}(\mathbf{X}_{1:L})\mathbf{W}_1$, where $\mathbf{W}_1 \in \mathbb{R}^{K \times D}$ is a learnable weight matrix.

Temporal Dependency Learner. The Temporal Dependency Learner is a core component of LuKAN, designed to capture temporal relationships within the motion sequence data. We employ a KAN layer, with associated matrix $\mathbf{\Phi} = (\phi_{q,p})$, where each trainable function $\phi_{q,p}$ is parameterized by a weighted linear combination of Lucas polynomials: $P_r(x) = xP_{r-1}(x) + P_{r-2}(x)$.

Residual Skip Connection. This skip connection links the input of KAN directly to its output, creating a residual pathway. Specifically, the output of the Temporal Dependency Learner is an $(L_a + L_d) \times D$ matrix given by $\mathbf{Z_2} = \mathrm{LN}(\mathrm{KAN}(\mathbf{Z_1})) + \mathbf{Z_1}$, where KAN and LN are applied along the temporal dimension.

Spatial Projection and IDWT. Together, the spatial projection and IDWT refine joint relationships and reconstruct the motion sequence in the time domain, resulting in an $L \times K$ output expressed as $\mathbf{Z}_3 = \mathrm{IDWT}(\mathbf{Z}_2\mathbf{W}_2)$, where $\mathbf{W}_2 \in \mathbb{R}^{D \times K}$ is a learnable matrix.

Model Prediction. The predicted sequence is a $T \times K$ matrix given by $\hat{\mathbf{X}}_{L+1:L+T} = \widetilde{\mathbf{Z}}_3 + \mathbf{X}_L$, where T is the prediction horizon, $\widetilde{\mathbf{Z}}_3$ consists of the first T rows of $\mathbf{Z}_3$, and $\mathbf{X}_L \in \mathbb{R}^{T \times K}$ is constructed by replicating the final pose $\mathbf{x}_L$ of the historical motion sequence T times.

Model Training. We train our model using the following loss function

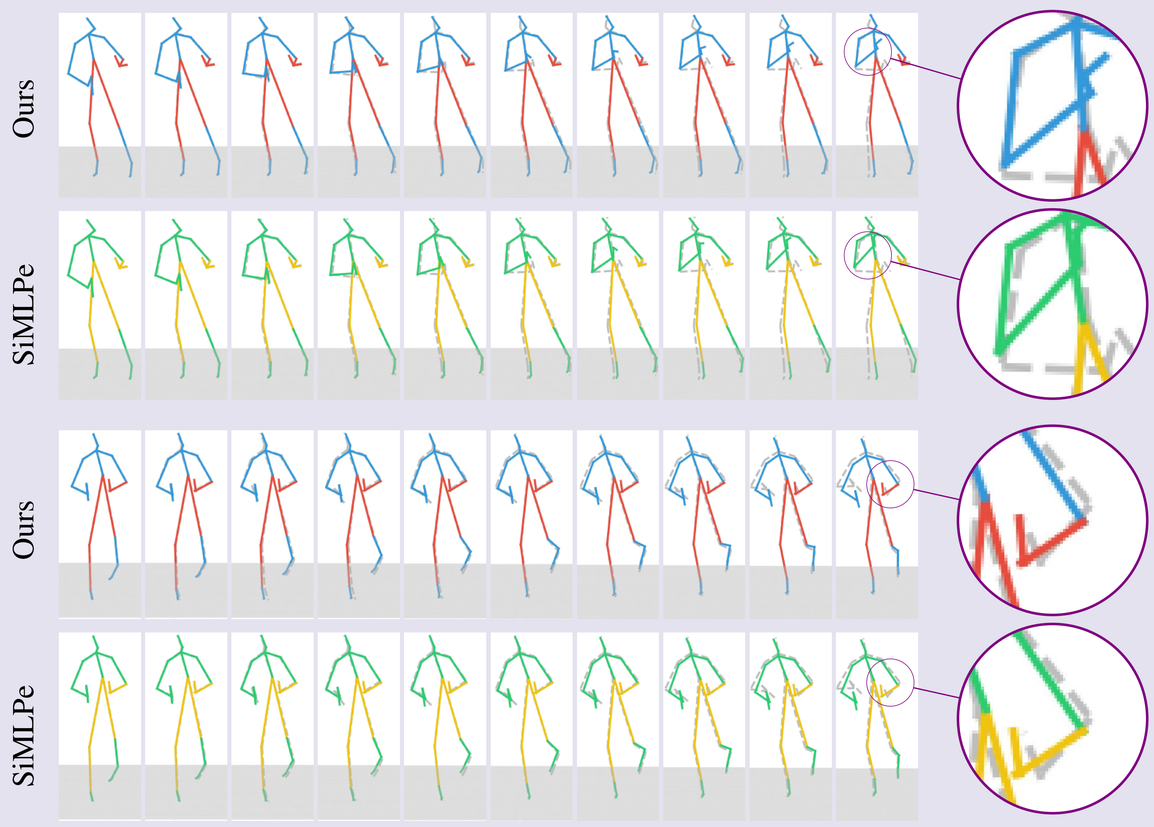
$$\mathcal{L} = \frac{1}{T} \sum_{t=L+1}^{L+T} (\|\mathbf{x}_t - \hat{\mathbf{x}}_t\|_2 + \|\mathbf{v}_t - \hat{\mathbf{v}}_t\|_2).$$

Results

Average MPJPE results of our model and baseline methods on Human3.6M for different prediction time steps in milliseconds (ms) ranging from 80ms to 1000ms.

	MPJPE (mm)↓							
	80	160	320	400	560	720	880	1000
ConvSeq2Seq	16.6	33.3	61.4	72.7	90.7	104.7	116.7	124.2
LTD-10-10	11.2	23.4	47.9	58.9	78.3	93.3	106.0	114.0
Hisrep	10.4	22.6	47.1	58.3	77.3	91.8	104.1	112.1
DMGNN	17.0	33.6	65.9	79.7	103	-	-	137.2
MSR-GCN	11.3	24.3	50.8	61.9	80.0	-	-	112.9
ST-DGCN	10.6	23.1	47.1	57.9	<u>76.3</u>	90.7	102.4	109.7
SPGSN	10.4	22.3	47	58.2	77.4	-	-	109.6
CIST-GCN	10.5	23.2	47.9	59.0	77.2	-	-	110.3
MotionMixer	11	23.6	47.8	59.3	77.8	91.4	106	111
SiMLPe	<u>9.6</u>	<u>21.7</u>	<u>46.3</u>	<u>57.3</u>	**75.7**	<u>90.1</u>	<u>101.8</u>	<u>109.4</u>
LuKAN (ours)	**9.4**	**21.5**	**46.2**	**57.2**	**75.7**	**89.9**	**101.6**	**109.3**

Visual comparison results on two actions: Directions (top) and Eating (bottom).



Ablations Studies:

		MPJPE (mm)↓							
		80	160	320	400	560	720	880	1000
Human3.6M	DCT	**9.4**	**21.4**	**45.8**	**56.8**	**75.7**	90.2	101.8	109.5
	DWT	**9.4**	21.5	46.2	57.2	**75.7**	**89.9**	**101.6**	**109.3**
AMASS	DCT	10.9	19.7	34.9	41.5	51.6	58.7	63.6	66.9
	DWT	**10.6**	**19.3**	**34.4**	**40.8**	**50.9**	**57.6**	**62.7**	**66.4**
3DPW	DCT	12.2	22.2	38.4	44.9	55.1	62.3	**68.1**	**72.2**
	DWT	**11.9**	**21.8**	**37.9**	**44.4**	**54.9**	**62.2**	**68.1**	**72.2**

Polynomials	MPJPE (mm)↓							
	80	160	320	400	560	720	880	1000
B-Splines	10.3	23.3	49.0	60.1	78.7	92.7	104.5	112.1
Chebyshev	9.7	22.1	47.2	58.3	77.1	91.7	103.7	111.6
Legendre	9.6	21.8	46.8	57.9	76.3	90.4	102.4	110.1
Hermite	9.5	21.6	46.3	57.3	76.0	90.3	102.2	110.1
Lucas	**9.4**	**21.5**	**46.2**	**57.2**	**75.7**	**89.9**	**101.6**	**109.3**