ProtoPairNet: Interpretable Regression through Prototypical Pair Reasoning

Rose Gurung^a, Ronilo Ragodos^b, Chiyu Ma^c, Tong Wang^d, Chaofan Chen^a

- ^a University of Maine^b University of New Hampshire
- ^c Dartmouth College ^d Yale University



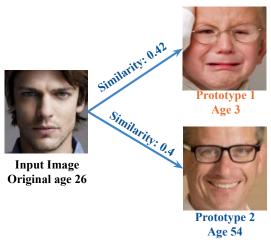
Input Image Original age 26

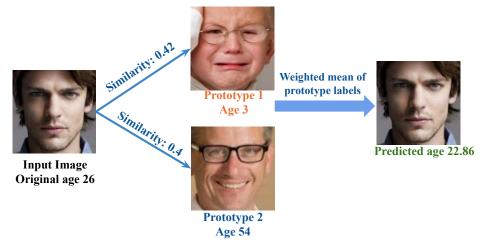


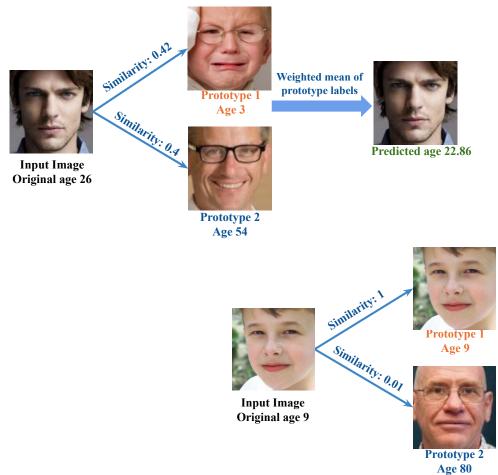
Prototype 1
Age 3

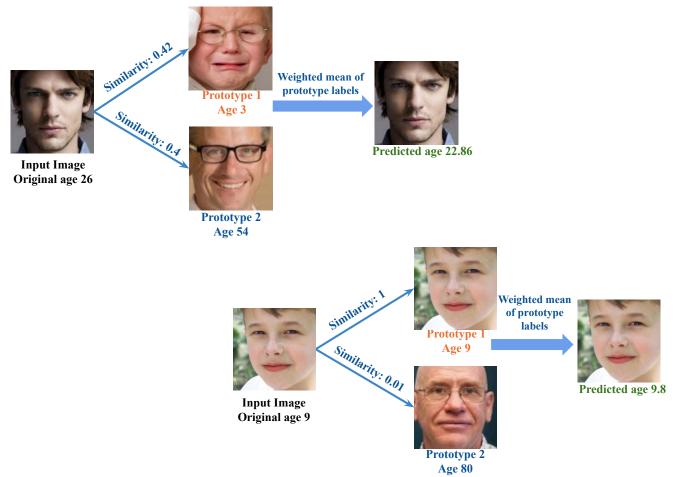


Prototype 2 Age 54







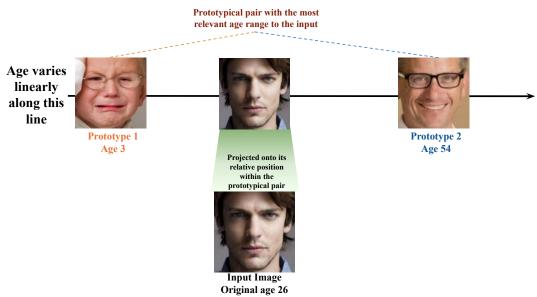


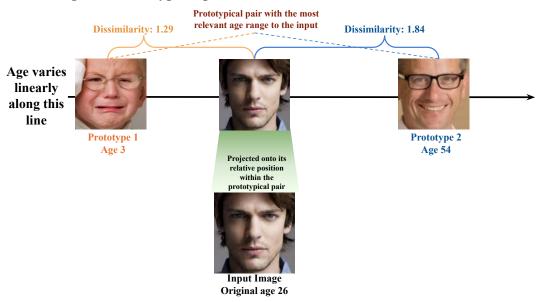


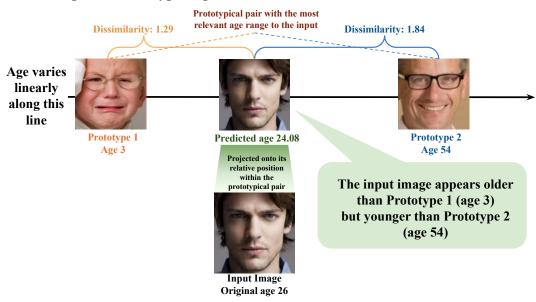


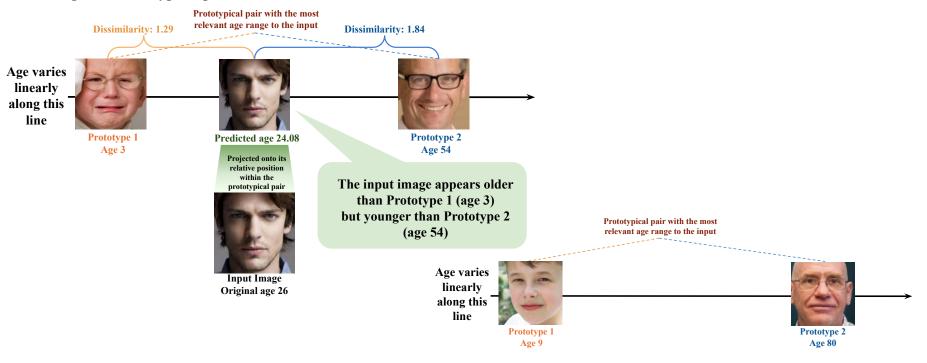


Input Image Original age 26

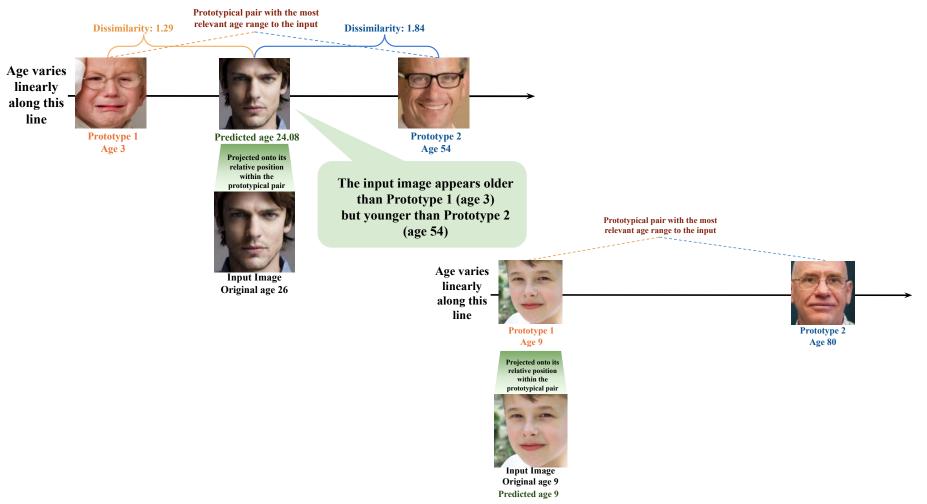


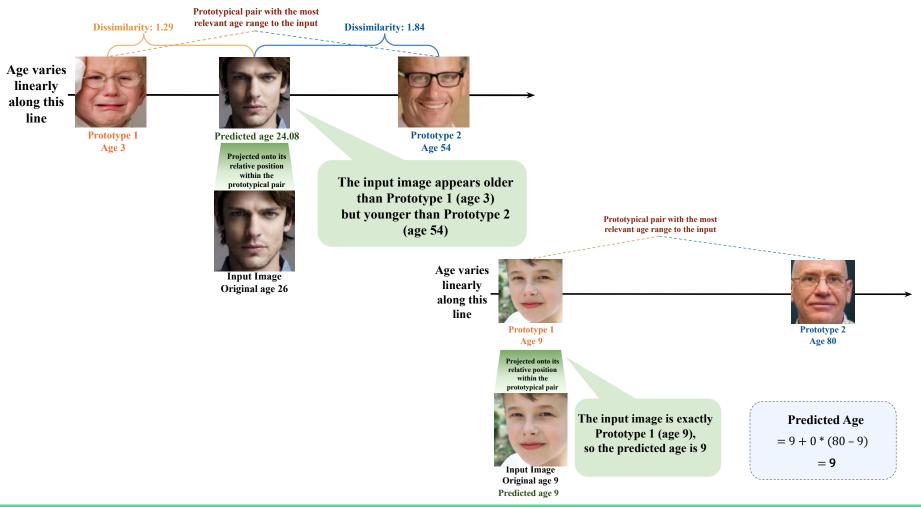


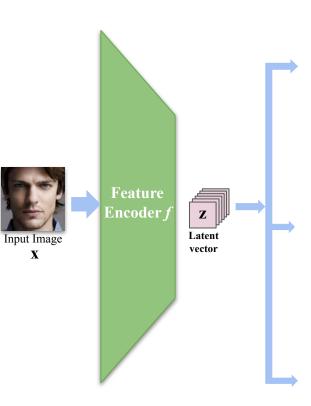












Prototypical pairs and corresponding labels: $(\mathbf{p}_{1}\;,\;\mathbf{p}_{2}\;),...,(\;\mathbf{p}_{2k-1}\;,\;\mathbf{p}_{2k}),...,(\;\mathbf{p}_{m-1}\;,\;\mathbf{p}_{m}),\textit{and}\;(\;y_{\mathbf{p}_{1}}\;,y_{\mathbf{p}_{2}}),...,(\;y_{\mathbf{p}_{2k-1}}\;,y_{\mathbf{p}_{2k}}),...,(\;y_{\mathbf{p}_{m-1}}\;,\;y_{\mathbf{p}_{m}})$ Pair p₁, p₂ $\mathbf{p_1}, y_{\mathbf{p_1}}$ $\mathbf{p}_2, y_{\mathbf{p}_2}$ $d_{(1,2)}$ Pair \mathbf{p}_{2k-1} , \mathbf{p}_{2k} $\mathbf{p}_{2k}, y_{\mathbf{p}_{2k}}$ $\mathbf{p}_{2k-1}, y_{\mathbf{p}_{2k-1}}$ Feature Encoder f Input projected onto the pair $d_{(2k-1, 2k)}$ Latent Input Image vector X Pair \mathbf{p}_{m-1} , \mathbf{p}_m $\mathbf{p}_{m-1}, y_{\mathbf{p}_{m-1}} \ \mathbf{p}_{m}, y_{\mathbf{p}_{m}}$

 $d_{(m-1, m)}$

Prototypical pairs and corresponding labels: $(\mathbf{p}_{1}\;,\;\mathbf{p}_{2}\;),\,...,(\;\mathbf{p}_{2k-1}\;,\;\mathbf{p}_{2k}),\,...,(\;\mathbf{p}_{m-1}\;,\;\mathbf{p}_{m}),\,\textit{and}\,(\;y_{\mathbf{p}_{1}}\;,y_{\mathbf{p}_{2}}),\,...,(\;y_{\mathbf{p}_{2k-1}}\;,y_{\mathbf{p}_{2k}}),\,...,(\;y_{\mathbf{p}_{m-1}}\;,\;y_{\mathbf{p}_{m}})$ Pair p₁, p₂ $\mathbf{p_1}, y_{\mathbf{p_1}}$ $\mathbf{p}_2, y_{\mathbf{p}_2}$ **Generating predictions** using equation (4) $\hat{y}_{(1,2)}$ $d_{(1,2)}$ Input projected onto the pair $\hat{\mathcal{Y}}_{(2k-1, 2k)}$ **Prediction** branch Pair \mathbf{p}_{2k-1} , \mathbf{p}_{2k} $\hat{\mathcal{Y}}_{(m-1, m)}$ $\mathbf{p}_{2k}, y_{\mathbf{p}_{2k}}$ $\mathbf{p}_{2k-1}, y_{\mathbf{p}_{2k-1}}$ **Feature** Encoder f Input projected onto the pair $d_{(2k-1, 2k)}$ Latent Input Image vector X Pair \mathbf{p}_{m-1} , \mathbf{p}_m $\mathbf{p}_{m-1}, y_{\mathbf{p}_{m-1}} \ \mathbf{p}_{m}, y_{\mathbf{p}_{m}}$ Input projected onto the pair line $d_{(m-1, m)}$

Prototypical pairs and corresponding labels: $(\mathbf{p}_1, \mathbf{p}_2), ..., (\mathbf{p}_{2k-1}, \mathbf{p}_{2k}), ..., (\mathbf{p}_{m-1}, \mathbf{p}_m),$ and $(y_{\mathbf{p}_1}, y_{\mathbf{p}_2}), ..., (y_{\mathbf{p}_{2k-1}}, y_{\mathbf{p}_{2k}}), ..., (y_{\mathbf{p}_{m-1}}, y_{\mathbf{p}_m})$ Pair p₁, p₂ $\mathbf{p_1}, y_{\mathbf{p_1}}$ $\mathbf{p}_2, y_{\mathbf{p}_2}$ **Generating predictions** using equation (4) $\hat{y}_{(1,2)}$ $d_{(1,2)}$ Input projected onto the pair $\hat{\mathcal{Y}}_{(2k-1, 2k)}$ **Prediction** branch Pair \mathbf{p}_{2k-1} , \mathbf{p}_{2k} $\mathbf{p}_{2k}, y_{\mathbf{p}_{2k}}$ $\hat{\mathcal{Y}}_{(m-1, m)}$ $\mathbf{p}_{2k-1}, y_{\mathbf{p}_{2k-1}}$ Feature Encoder j Input projected onto the pair $d_{(2k-1, 2k)}$ Latent Input Image vector \mathbf{X} $r_{(1, 2)}$ Pair \mathbf{p}_{m-1} , \mathbf{p}_m Pair relevance $r_{(2k-1, 2k)}$ branch $\mathbf{p}_{m-1}, y_{\mathbf{p}_{m-1}} \ \mathbf{p}_{m}, y_{\mathbf{p}_{m}}$ Input projected onto the pair line $d_{(m-1, m)}$ $r_{(m-1, m)}$ **Computing relevance scores** using equation (5)

Prototypical pairs and corresponding labels: $(\mathbf{p}_1, \mathbf{p}_2), ..., (\mathbf{p}_{2k-1}, \mathbf{p}_{2k}), ..., (\mathbf{p}_{m-1}, \mathbf{p}_m),$ and $(y_{\mathbf{p}_1}, y_{\mathbf{p}_2}), ..., (y_{\mathbf{p}_{2k-1}}, y_{\mathbf{p}_{2k}}), ..., (y_{\mathbf{p}_{m-1}}, y_{\mathbf{p}_m})$ Pair $\mathbf{p}_1, \mathbf{p}_2$ $\mathbf{p_1}, y_{\mathbf{p_1}}$ $\mathbf{p}_2, y_{\mathbf{p}_2}$ **Generating predictions** using equation (4) $\hat{y}_{(1,2)}$ $d_{(1,2)}$ Input projected onto the pair $\hat{\mathcal{Y}}_{(2k-1, 2k)}$ **Prediction** branch Pair \mathbf{p}_{2k-1} , \mathbf{p}_{2k} Predicted age $\mathbf{p}_{2k}, y_{\mathbf{p}_{2k}}$ $\hat{\mathcal{Y}}_{(m-1, m)}$ $\mathbf{p}_{2k-1}, y_{\mathbf{p}_{2k-1}}$ **Final** 24.08 Using prediction Feature equation (6) Encoder f X Input projected onto the pair $d_{(2k-1, 2k)}$ Latent Input Image vector \mathbf{X} $r_{(1, 2)}$ Pair \mathbf{p}_{m-1} , \mathbf{p}_m Pair relevance $r_{(2k-1, 2k)}$ branch $\mathbf{p}_{m-1}, y_{\mathbf{p}_{m-1}} \ \mathbf{p}_{m}, y_{\mathbf{p}_{m}}$ Input projected onto the pair $d_{(m-1, m)}$ $r_{(m-1, m)}$ **Computing relevance scores** using equation (5)

Table 1: Age prediction results. Comparision of ProtoPairNet with baseline black-box models (without prototypes) in terms of MAE and R^2 scores across different architectures.

Architecture	ResNet50		EfficientNet-B0		VGG19		ViT-Small (Patch 16)	
	MAE	R ² Score	MAE	R ² Score	MAE	R ² Score	MAE	R ² Score
Baseline	4.71 ± 0.09	0.86 ± 0.004	4.82 ± 0.08	0.86 ± 0.004	4.73 ± 0.07	0.86 ± 0.005	4.81 ± 0.06	0.86 ± 0.004
ProtoPairNet (Ours)	$\textbf{4.59} \pm \textbf{0.01}$	$\textbf{0.87} \pm \textbf{0.001}$	4.88 ± 0.03	0.86 ± 0.0006	$\textbf{4.57} \pm \textbf{0.01}$	$\textbf{0.87} \pm \textbf{0.001}$	$\textbf{4.63} \pm \textbf{0.04}$	$\textbf{0.87} \pm \textbf{0.001}$

Table 2: Comparison of ProtoPairNet with other prototype-based ResNet50 models for regression using MAE and \mathbb{R}^2 scores.

Model	MAE	\mathbf{R}^2 score
HPN	5.08 ± 0.03	0.84 ± 0.001
INSightR-Net (1×1)	14.01 ± 0.01	-0.03 ± 0.001
INSightR-Net (full)	14.81 ± 0.40	-0.08 ± 0.07
ExPeRT	17.36 ± 2.79	-0.37 ± 0.41
ProtoPairNet (Ours)	$\textbf{4.59} \pm \textbf{0.01}$	$\textbf{0.87} \pm \textbf{0.001}$

Table 3: Car Racing results. Comparison of different architectures in terms of rewards.

Architecture	Reward		
k-Means[21]	-2.09 ± 0.94		
PW-Net*[21]	-9.48 ± 2.50		
PW-Net[21]	220.61 ± 0.70		
ProtoPairNet	223.97 ± 0.75		
	gent 221.36 ± 0.96		

ProtoPairNet: Interpretable Regression through Prototypical Pair Reasoning

Rose Gurung

University of Maine rose.gurung@maine.edu

Ronilo Ragodos

University of New Hampshire Ronilo.Ragodos@unh.edu

Chiyu Ma

Dartmouth College chiyu.ma.gr@dartmouth.edu

Tong Wang

Yale University tong.wang.tw687@yale.edu

Chaofan Chen

University of Maine chaofan.chen@maine.edu