Table S1. To evaluate the method's ability to more challenging situations through large different data distributions, we used three general pancreas datasets which are different in species and platform. So, we applied both the Baron mouse and the Tabula Muris mouse cell datasets as source domain data, and the Segerstolpe human cell dataset as target domain data. We used the SingleCellNet method (https://github.com/pcahan1/singleCellNet) for converting the names for genes in the cross-species part and obtained common genes among the mouse and human datasets. The color spectrum from light blue to dark blue shows an increase in accuracy.

Source → Target	Classification accuracy (%) on each class in pancreas datasets								Mean Accuracy (%)
	Alpha	Beta	Acinar	Gamma	Ductal	Endothelial	Delta	Stellate	
Baron mouse dataset & Tabula Muris mouse dataset → Segerstolpe human dataset	0.98	0.98	0.97	1.00	99	1.00	1.00	0.98	**0.99**

Table S2. The ocular datasets used in this study. The number of cells in each subtype is shown. We have used the cell type annotations based on original papers. There are a few transcriptional correspondences between subtypes of P5 RGCs (Rheaume et al.) and adult RGC types (Tran et al.), which we used based on Tran's study reported.

Tran et al.		Rheaume et al.		Macosko et al.	
# Cells	Sub types	# Cells	Sub types	# Cells	Sub types
3,000	W3-like1: Adult RGC1	161	P5 RGC0	252	Horizontal cells
2,859	W3D1: Adult RGC2	426	W3D1: P5 RGC1	432	Retinal ganglion cells
1,990	F-mini-ON: Adult RGC3	188	P5 RGC2	289	Amacrine cells
1,868	F-mini-OFF: Adult RGC4	196	P5 RGC3	73	Amacrine cells
1,715	J-RGC: Adult RGC5	329	F-mini-ON: P5 RGC4	77	Amacrine cells
1,590	W3B: Adult RGC6	66	P5 RGC5	211	Amacrine cells
1,579	Adult RGC7	52	P5 RGC6	326	Amacrine cells
1,258	Adult RGC8	268	F-mini-OFF: P5 RGC7	159	Amacrine cells
1,223	T-RGC: Adult RGC9	95	F-RGC: P5 RGC8	350	Amacrine cells
1,170	Adult RGC10	185	P5 RGC9	191	Amacrine cells
990	Adult RGC11	143	P5 RGC10	214	Amacrine cells
953	ooDSGC-N: Adult RGC12	88	P5 RGC11	274	Amacrine cells
943	W3-like2: Adult RGC13	135	P5 RGC12	50	Amacrine cells
875	Adult RGC14	429	W3-like1: P5 RGC13	111	Amacrine cells
865	Adult RGC15	235	J-RGC: P5 RGC14	73	Amacrine cells
829	ooDSGC-D/V: Adult RGC16	91	T-RGC-S2: P5 RGC15	262	Amacrine cells
828	T-RC-S1: Adult RGC17	135	P5 RGC16	375	Amacrine cells
826	Adult RGC18	121	P5 RGC17	83	Amacrine cells
775	Adult RGC19	80	W3D3: P5 RGC18	127	Amacrine cells
711	Adult RGC20	115	P5 RGC19	389	Amacrine cells
687	T-RGC-S2: Adult RGC21	224	P5 RGC20	254	Amacrine cells
610	MX: Adult RGC22	102	P5 RGC21	274	Amacrine cells
601	W3D2: Adult RGC23	48	P5 RGC22	264	Amacrine cells
553	Adult RGC24	93	P5 RGC23	29,400	Rods
542	Adult RGC25	89	P5 RGC24	1,868	Cones
534	Adult RGC26	175	P5 RGC25	2,217	Bipolar cells
529	Adult RGC27	233	P5 RGC26	664	Bipolar cells
517	F-midi-OFF: Adult RGC28	147	W3-like2: P5 RGC27	496	Bipolar cells
499	Adult RGC29	100	P5 RGC28	591	Bipolar cells
491	W3D3: Adult RGC30	124	P5 RGC29	636	Bipolar cells
444	M2: Adult RGC31	186	P5 RGC30	512	Bipolar cells
407	F-RGC: Adult RGC32	168	ooDSGC-N: P5 RGC31	320	Bipolar cells
323	M1a: Adult RGC33	133	P5 RGC32	849	Bipolar cells
312	Adult RGC34	108	P5 RGC33	1,624	Muller glia
310	Adult RGC35	155	W3B: P5 RGC34	54	Astrocytes
236	Adult RGC36	70	P5 RGC35	85	Fibroblasts
213	Adult RGC37	183	P5 RGC36	252	Vascular endothelium
207	F-midi-ON: Adult RGC38	135	P5 RGC37	63	Pericytes
202	Adult RGC39	150	P5 RGC38	67	Microglia
174	M1b: Adult RGC40	44	P5 RGC39		
126	aON-T: Adult RGC41	20	P5 RGC40		
113	aOFF-S: Adult RGC42				
106	aON-S/M4: Adult RGC43				
62	Adult RGC44				
54	aOFF-T: Adult RGC45				

Figure S1. Subsampling cells and genes in the first and second datasets to assess the impact on the model's accuracy. Panel A shows the accuracy versus different number of selected cells when the number of genes is unchanged. Panel B shows the accuracy of the model versus different number of randomly selected genes when the number of cells is unchanged. Selecting both lower number of cells and genes would compromise the accuracy of the model.

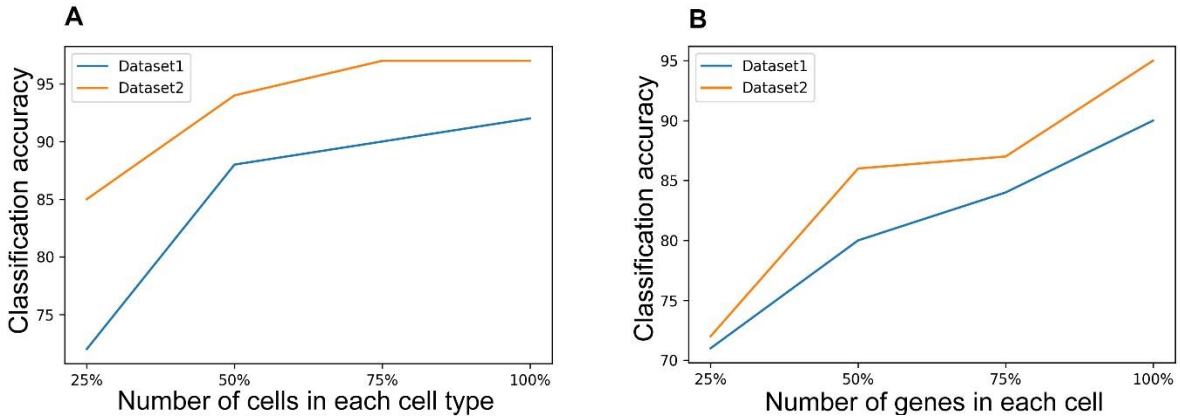

Figure S2. Selecting different numbers of highly variable genes to investigate the impact on the model's accuracy. While selecting over 2000 top highly variable genes would not improve the accuracy, selecting significantly lower number of top highly variable genes would compromise the accuracy.