Digit	True positives (low)	True positives (high)	False negatives (low)	False negatives (high)	True negatives (low)	True negatives (high)	False positives (low)	False positives (high)	Precision	Recall	F1 Score	total
0	427	7	6	6	8	1765	4	5	0.9797	0.9731	0.9764	2228
1	421	10	4	11	15	1761	2	4	0.9863	0.9664	0.9762	2228
2	412	19	12	3	143	1627	7	5	0.9729	0.9664	0.9696	2228
3	403	18	17	8	138	1628	7	9	0.9634	0.9439	0.9536	2228
4	388	30	15	13	44	1725	7	6	0.9698	0.9372	0.9532	2228
5	393	31	14	8	33	1739	5	5	0.9770	0.9507	0.9636	2228
6	411	18	7	10	18	1762	1	1	0.9954	0.9619	0.9783	2228
7	396	21	18	11	52	1722	4	4	0.9812	0.9350	0.9575	2228
8	371	33	18	24	44	1718	8	12	0.9528	0.9058	0.9287	2228
9	394	24	10	18	91	1668	10	13	0.9478	0.9372	0.9425	2228
Mean	401.6	21.1	12.1	11.2	58.6	1711.5	5.5	6.4	0.9726	0.9478	0.9600	2228
											0.0162	

Training and testing with transformations

Digit	True positives (low)	True positives (high)	False negatives (low)	False negatives (high)	True negatives (low)	True negatives (high)	False positives (low)	False positives (high)	Precision	Recall	F1 Score	total
0	434	7	2	3	6	1771	3	2	0.9888	0.9888	0.9888	2228
1	425	8	8	5	12	1765	0	5	0.9886	0.9709	0.9796	2228
2	422	10	11	3	117	1661	1	3	0.9908	0.9686	0.9796	2228
3	418	10	12	6	115	1661	2	4	0.9862	0.9596	0.9727	2228
4	414	14	7	11	22	1750	5	5	0.9772	0.9596	0.9683	2228
5	420	12	6	8	16	1755	3	8	0.9752	0.9686	0.9719	2228
6	421	8	8	9	7	1772	3	0	0.9931	0.9619	0.9772	2228
7	417	12	14	3	21	1751	5	5	0.9772	0.9619	0.9695	2228
8	383	28	21	14	26	1745	3	8	0.9739	0.9215	0.9470	2228
9	400	20	16	10	53	1721	5	3	0.9813	0.9417	0.9611	2228
Mean	415.4	12.9	10.5	7.2	39.5	1735.2	3	4.3	0.9832	0.9603	0.9716	2228
											0.0115	

Training with transformations, testing without transformations

Digit	True positives (low)	True positives (high)	False negatives (low)	False negatives (high)	True negatives (low)	True negatives (high)	False positives (low)	False positives (high)	Precision	Recall	F1 Score	total
0	374	25	11	36	13	1755	7	7	0.9661	0.8946	0.9290	2228
1	337	32	38	39	40	1716	14	12	0.9342	0.8274	0.8775	2228
2	354	35	23	34	29	1722	15	16	0.9262	0.8722	0.8984	2228
3	369	25	13	39	30	1726	14	12	0.9381	0.8834	0.9099	2228
4	364	26	23	33	27	1722	21	12	0.9220	0.8744	0.8976	2228
5	373	18	15	40	28	1735	12	7	0.9537	0.8767	0.9136	2228
6	369	22	18	37	34	1723	14	11	0.9399	0.8767	0.9072	2228
7	359	26	24	37	42	1713	11	16	0.9345	0.8632	0.8974	2228
8	355	27	26	38	64	1668	21	29	0.8843	0.8565	0.8702	2228
9	315	49	28	54	53	1674	32	23	0.8687	0.8161	0.8416	2228
Mean	356.9	28.5	21.9	38.7	36	1715.4	16.1	14.5	0.9268	0.8641	0.8942	2228
											0.0251	

Training without transformations, testing with transformations

Digit	True positives (low)	True positives (high)	False negatives (low)	False negatives (high)	True negatives (low)	True negatives (high)	False positives (low)	False positives (high)	Precision	Recall	F1 Score	total
0	435	3	2	6	3	1775	3	1	0.9910	0.9821	0.9865	2228
1	433	2	0	11	4	1776	0	2	0.9954	0.9753	0.9853	2228
2	404	14	12	16	9	1766	3	4	0.9835	0.9372	0.9598	2228
3	417	9	10	10	13	1756	6	7	0.9704	0.9552	0.9627	2228
4	417	8	9	12	4	1769	8	1	0.9793	0.9529	0.9659	2228
5	410	13	10	13	11	1762	4	5	0.9792	0.9484	0.9636	2228
6	417	8	1	20	1	1774	4	3	0.9838	0.9529	0.9681	2228
7	406	14	12	14	13	1761	5	3	0.9813	0.9417	0.9611	2228
8	394	18	14	20	24	1744	6	8	0.9671	0.9238	0.9450	2228
9	388	22	15	21	16	1746	10	10	0.9535	0.9193	0.9361	2228
Mean	412.1	11.1	8.5	14.3	9.8	1762.9	4.9	4.4	0.9784	0.9489	0.9634	2228
											0.0154	

Training without transformations, testing without transformations