









Table 1: Summary of Results

Group [†]			Metric [‡]		
Norm	Func	Loss	Dice Score	Sensitivity	Positive Predictive Value
BatchNorm	LeakyReLU	BCE	0.54, 0.72 (IQR 0.08-0.86)	0.87, 0.88 (IQR 0.86-0.90)	0.68, 0.69 (IQR 0.64-0.72)
		Focal	0.55, 0.74 (IQR 0.08-0.87)	0.87, 0.88 (IQR 0.85-0.89)	0.73, 0.74 (IQR 0.70-0.77)
	ReLU	BCE	0.51, 0.77 (IQR 0.00-0.89)	0.85, 0.86 (IQR 0.83-0.87)	0.75, 0.76 (IQR 0.72-0.78)
		Focal	0.61, 0.82 (IQR 0.24-0.90)	0.87, 0.87 (IQR 0.85-0.89)	0.79, 0.79 (IQR 0.77-0.82)
LayerNorm	LeakyReLU	BCE	0.66, 0.83 (IQR 0.61-0.90)	0.86, 0.87 (IQR 0.84-0.88)	0.84, 0.84 (IQR 0.82-0.86)
		Focal	0.64, 0.85 (IQR 0.28-0.90)	0.83, 0.84 (IQR 0.80-0.86)	0.87, 0.87 (IQR 0.85-0.88)
	ReLU	BCE	0.64, 0.86 (IQR 0.42-0.90)	0.85, 0.86 (IQR 0.83-0.88)	0.85, 0.85 (IQR 0.84-0.86)
		Focal	0.65, 0.84 (IQR 0.46-0.91)	0.86, 0.86 (IQR 0.84-0.88)	0.87, 0.87 (IQR 0.86-0.88)

[†] Results reported for: valid cohort, tumor = positive, repeats = 1-2, filters = lin-16, class weights = 1, resnet = False, convnext = False

[‡] Table values are shown as: mean, median (interquartile range)









Table 2: Comparison Between Experiments - Dice Score

Group [†]				Violin	Exp [‡]							
Norm	Func	Loss	Exp	Dice Score	1	2	3	4	5	6	7	8
BatchNorm	LeakyReLU	BCE	1		–	<i>N.S.</i>	<i>N.S.</i>	*	**	*	*	**
		Focal	2		<i>N.S.</i>	–	<i>N.S.</i>	*	*	<i>N.S.</i>	<i>N.S.</i>	*
	ReLU	BCE	3		<i>N.S.</i>	<i>N.S.</i>	–	*	**	*	*	*
		Focal	4		*	*	*	–	<i>N.S.</i>	<i>N.S.</i>	<i>N.S.</i>	<i>N.S.</i>
LayerNorm	LeakyReLU	BCE	5		**	*	**	<i>N.S.</i>	–	<i>N.S.</i>	<i>N.S.</i>	<i>N.S.</i>
		Focal	6		*	<i>N.S.</i>	*	<i>N.S.</i>	<i>N.S.</i>	–	<i>N.S.</i>	<i>N.S.</i>
	ReLU	BCE	7		*	<i>N.S.</i>	*	<i>N.S.</i>	<i>N.S.</i>	<i>N.S.</i>	–	<i>N.S.</i>
		Focal	8		**	*	*	<i>N.S.</i>	<i>N.S.</i>	<i>N.S.</i>	<i>N.S.</i>	–

[†] Results reported for: valid cohort, tumor = positive, repeats = 1-2, filters = lin-16, class weights = 1, resnet = False, convnext = False

[‡] *N.S.* = not significant, * = $p < 0.05$, ** = $p < 0.005$


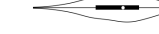






Table 3: Comparison Between Experiments - Sensitivity

Group [†]				Violin	Exp [‡]							
Norm	Func	Loss	Exp	Sensitivity	1	2	3	4	5	6	7	8
BatchNorm	LeakyReLU	BCE	1		–	*	**	<i>N.S.</i>	**	**	**	**
		Focal	2		*	–	**	<i>N.S.</i>	**	**	**	**
	ReLU	BCE	3		**	**	–	**	**	**	<i>N.S.</i>	**
		Focal	4		<i>N.S.</i>	<i>N.S.</i>	**	–	**	**	**	**
LayerNorm	LeakyReLU	BCE	5		**	**	**	**	–	**	**	<i>N.S.</i>
		Focal	6		**	**	**	**	**	–	**	**
	ReLU	BCE	7		**	**	<i>N.S.</i>	**	**	**	–	**
		Focal	8		**	**	**	**	<i>N.S.</i>	**	**	–

[†] Results reported for: valid cohort, tumor = positive, repeats = 1-2, filters = lin-16, class weights = 1, resnet = False, convnext = False

[‡] *N.S.* = not significant, * = $p < 0.05$, ** = $p < 0.005$

Table 4: Comparison Between Experiments - Positive Predictive Value

Group [†]				Violin	Exp [‡]							
Norm	Func	Loss	Exp	Positive Predictive Value	1	2	3	4	5	6	7	8
BatchNorm	LeakyReLU	BCE	1		—	**	**	**	**	**	**	**
		Focal	2		**	—	**	**	**	**	**	**
	ReLU	BCE	3		**	**	—	**	**	**	**	**
		Focal	4		**	**	**	—	**	**	**	**
LayerNorm	LeakyReLU	BCE	5		**	**	**	**	—	**	**	**
		Focal	6		**	**	**	**	**	—	**	**
	ReLU	BCE	7		**	**	**	**	**	**	—	**
		Focal	8		**	**	**	**	**	**	**	—

[†] Results reported for: valid cohort, tumor = positive, repeats = 1-2, filters = lin-16, class weights = 1, resnet = False, convnext = False

[‡] ** = $p < 0.005$

Figure 1: Comparison Between Experiments - Violin Plots

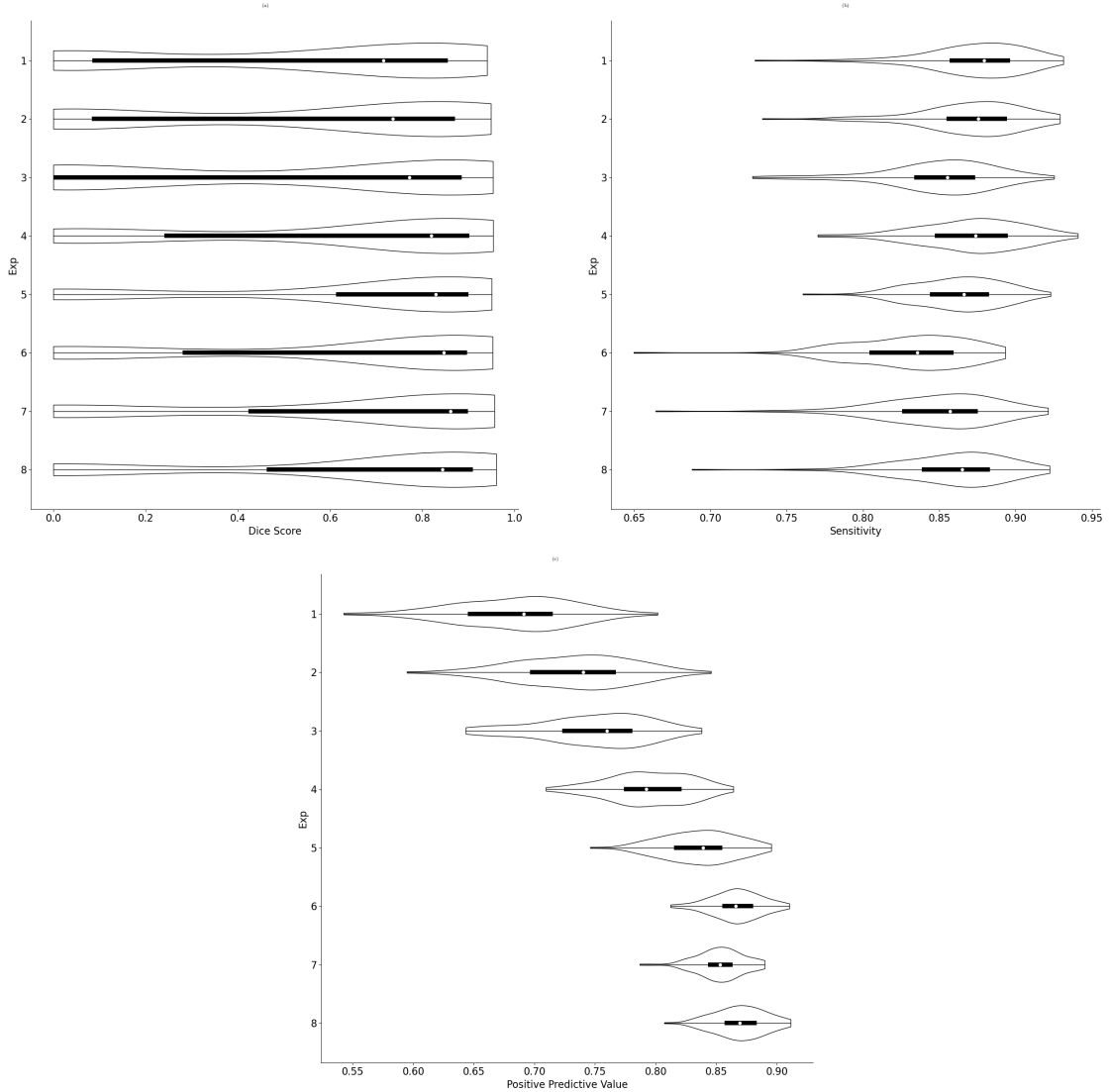


Table 5: Mean Differences - Dice Score

Group [†]				Exp							
Norm	Func	Loss	Exp	1	2	3	4	5	6	7	8
BatchNorm	LeakyReLU	BCE	1	–	0.01 (p = 0.795)	-0.03 (p = 0.457)	0.07 (p = 0.025)	0.12 (p < 0.001)	0.10 (p = 0.007)	0.10 (p = 0.005)	0.11 (p = 0.002)
		Focal	2	-0.01 (p = 0.795)	–	-0.04 (p = 0.289)	0.06 (p = 0.047)	0.11 (p = 0.013)	0.09 (p = 0.079)	0.09 (p = 0.061)	0.10 (p = 0.042)
	ReLU	BCE	3	0.03 (p = 0.457)	0.04 (p = 0.289)	–	0.10 (p = 0.016)	0.14 (p = 0.003)	0.12 (p = 0.020)	0.12 (p = 0.016)	0.13 (p = 0.009)
		Focal	4	-0.07 (p = 0.025)	-0.06 (p = 0.047)	-0.10 (p = 0.016)	–	0.05 (p = 0.227)	0.03 (p = 0.579)	0.03 (p = 0.572)	0.04 (p = 0.413)
	LayerNorm	LeakyReLU	BCE	5	-0.12 (p < 0.001)	-0.11 (p = 0.013)	-0.14 (p = 0.003)	-0.05 (p = 0.227)	– (p = 0.509)	-0.02 (p = 0.463)	-0.01 (p = 0.750)
			Focal	6	-0.10 (p = 0.007)	-0.09 (p = 0.079)	-0.12 (p = 0.020)	-0.03 (p = 0.579)	0.02 (p = 0.509)	0.00 (p = 0.960)	0.01 (p = 0.488)
		ReLU	BCE	7	-0.10 (p = 0.005)	-0.09 (p = 0.061)	-0.12 (p = 0.016)	-0.03 (p = 0.572)	0.02 (p = 0.463)	-0.00 (p = 0.960)	0.01 (p < 0.001)
			Focal	8	-0.11 (p = 0.002)	-0.10 (p = 0.042)	-0.13 (p = 0.009)	-0.04 (p = 0.413)	0.01 (p = 0.750)	-0.01 (p = 0.488)	– (p = 0.457)

[†] Results reported for: valid cohort, tumor = positive, repeats = 1-2, filters = lin-16, class weights = 1, resnet = False, convnext = False

Table 6: Mean Differences - Sensitivity

Group [†]				Exp							
Norm	Func	Loss	Exp	1	2	3	4	5	6	7	8
BatchNorm	LeakyReLU	BCE	1	–	-0.00 (p = 0.030)	-0.02 (p < 0.001)	-0.00 (p = 0.331)	-0.01 (p < 0.001)	-0.04 (p < 0.001)	-0.02 (p < 0.001)	-0.02 (p < 0.001)
		Focal	2	0.00 (p = 0.030)	–	-0.02 (p < 0.001)	0.00 (p = 0.997)	-0.01 (p = 0.002)	-0.04 (p < 0.001)	-0.02 (p < 0.001)	-0.01 (p < 0.001)
	ReLU	BCE	3	0.02 (p < 0.001)	0.02 (p < 0.001)	–	0.02 (p < 0.001)	0.01 (p < 0.001)	-0.02 (p < 0.001)	-0.00 (p = 0.953)	0.01 (p < 0.001)
		Focal	4	0.00 (p = 0.331)	-0.00 (p = 0.997)	-0.02 (p < 0.001)	–	-0.01 (p = 0.004)	-0.04 (p < 0.001)	-0.02 (p < 0.001)	-0.01 (p < 0.001)
	LayerNorm	LeakyReLU	BCE	5	0.01 (p < 0.001)	0.01 (p = 0.002)	-0.01 (p < 0.001)	0.01 (p = 0.004)	– (p < 0.001)	-0.03 (p < 0.001)	-0.01 (p < 0.001)
			Focal	6	0.04 (p < 0.001)	0.04 (p < 0.001)	0.02 (p < 0.001)	0.04 (p < 0.001)	0.03 (p < 0.001)	0.02 (p < 0.001)	0.03 (p < 0.001)
		ReLU	BCE	7	0.02 (p < 0.001)	0.02 (p < 0.001)	0.00 (p = 0.953)	0.02 (p < 0.001)	0.01 (p < 0.001)	-0.02 (p < 0.001)	0.01 (p < 0.001)
			Focal	8	0.02 (p < 0.001)	0.01 (p < 0.001)	-0.01 (p < 0.001)	0.01 (p < 0.001)	0.00 (p = 0.266)	-0.03 (p < 0.001)	– (p < 0.001)

[†] Results reported for: valid cohort, tumor = positive, repeats = 1-2, filters = lin-16, class weights = 1, resnet = False, convnext = False

Table 7: Mean Differences - Positive Predictive Value

Group [†]				Exp							
Norm	Func	Loss	Exp	1	2	3	4	5	6	7	8
BatchNorm	LeakyReLU	BCE	1	–	0.05 (p < 0.001)	0.06 (p < 0.001)	0.11 (p < 0.001)	0.15 (p < 0.001)	0.18 (p < 0.001)	0.17 (p < 0.001)	0.19 (p < 0.001)
		Focal	2	-0.05 (p < 0.001)	–	0.02 (p < 0.001)	0.06 (p < 0.001)	0.10 (p < 0.001)	0.13 (p < 0.001)	0.12 (p < 0.001)	0.14 (p < 0.001)
	ReLU	BCE	3	-0.06 (p < 0.001)	-0.02 (p < 0.001)	–	0.05 (p < 0.001)	0.09 (p < 0.001)	0.12 (p < 0.001)	0.10 (p < 0.001)	0.12 (p < 0.001)
		Focal	4	-0.11 (p < 0.001)	-0.06 (p < 0.001)	-0.05 (p < 0.001)	–	0.04 (p < 0.001)	0.07 (p < 0.001)	0.06 (p < 0.001)	0.08 (p < 0.001)
	LayerNorm	LeakyReLU	BCE	5	-0.15 (p < 0.001)	-0.10 (p < 0.001)	-0.09 (p < 0.001)	-0.04 (p < 0.001)	– (p < 0.001)	0.03 (p < 0.001)	0.02 (p < 0.001)
			Focal	6	-0.18 (p < 0.001)	-0.13 (p < 0.001)	-0.12 (p < 0.001)	-0.07 (p < 0.001)	-0.03 (p < 0.001)	– (p < 0.001)	0.00 (p < 0.001)
		ReLU	BCE	7	-0.17 (p < 0.001)	-0.12 (p < 0.001)	-0.10 (p < 0.001)	-0.06 (p < 0.001)	-0.02 (p < 0.001)	0.01 (p < 0.001)	0.02 (p < 0.001)
			Focal	8	-0.19 (p < 0.001)	-0.14 (p < 0.001)	-0.12 (p < 0.001)	-0.08 (p < 0.001)	-0.03 (p < 0.001)	-0.02 (p < 0.001)	– (p < 0.001)

[†] Results reported for: valid cohort, tumor = positive, repeats = 1-2, filters = lin-16, class weights = 1, resnet = False, convnext = False