

TO-DO

- ☒ Quantas vezes falha a inferência da expressão neutra nos testes?
- ☒ Percentagem de neutras e não neutras?
- ☒ Data Augmentation, em vez de diminuir número de imagens não neutras, que reduz dataset.
- ☒ Impacto do balanceamento do dataset?
- ☒ Impacto do pré-processamento?

Para o dataset JAFFE (sem balanceamento)

30 imagens Neutral (14%)

183 imagens Not Neutral (86%)

20% teste (43 imagens)

10% validação

70% treino

neutras classificadas como não neutras: 7 (16%)

não neutras classificadas como neutras: 0

neutras classificadas como neutras: 6

não neutras classificadas como não neutras: 30

Para o dataset FER2013 (sem balanceamento)

6198 imagens Neutral (17%)

29688 imagens Not Neutral (83%)

20% teste (7178)

10% validação

70% treino

neutras classificadas como não neutras: 895 (12%)

não neutras classificadas como neutras: 206

neutras classificadas como neutras: 324

não neutras classificadas como não neutras: 5753

Para o dataset FER2013 (balanceada)

6198 imagens Neutral (50%)

6198 imagens Not Neutral (50%)

20% teste (2480)

10% validação

70% treino

neutras classificadas como não neutras: 249 (10%)

não neutras classificadas como neutras: 439

neutras classificadas como neutras: 1016

não neutras classificadas como não neutras: 776

Para o dataset FER2013 (Data Augmentation)

Contraste: 0.5, 1.5 e 2.0 em todas as imagens (total = 24792 neutras, 46%)

29688 imagens Not Neutral (54%)

neutras classificadas como não neutras: 1018 (9%)

não neutras classificadas como neutras: 383

neutras classificadas como neutras: 3955

não neutras classificadas como não neutras: 5540

Para dataset CK+

593 neutras (64,5%)

327 não neutras (35,5%)

10% teste (92 imagens)

10% validação

80% treino

Para dataset CK+ balanceado

327 neutras (50%)

327 não neutras (50%)

10% teste (66 imagens)

10% validação

80% treino

Para dataset CK+ data augmentation

1186 neutras (54,73%)

981 não neutras (45,27%)

10% teste (217 imagens)

10% validação

80% treino

Model FER

Impacto do Pré Processamento

Stage 0: Raw Set

Test loss: 36,89%

Test accuracy: 83,88%

Train loss: 31,55%

Train accuracy: 85,51%

Validation loss: 39,86%

Validation accuracy: 81,72%

	precision	recall	f1-score	support
neutral	0.60	0.19	0.29	360
not neutral	0.85	0.97	0.91	1731
accuracy			0.84	2091
macro avg	0.73	0.58	0.60	2091
weighted avg	0.81	0.84	0.80	2091

[[70 290] acertou em 83,9%

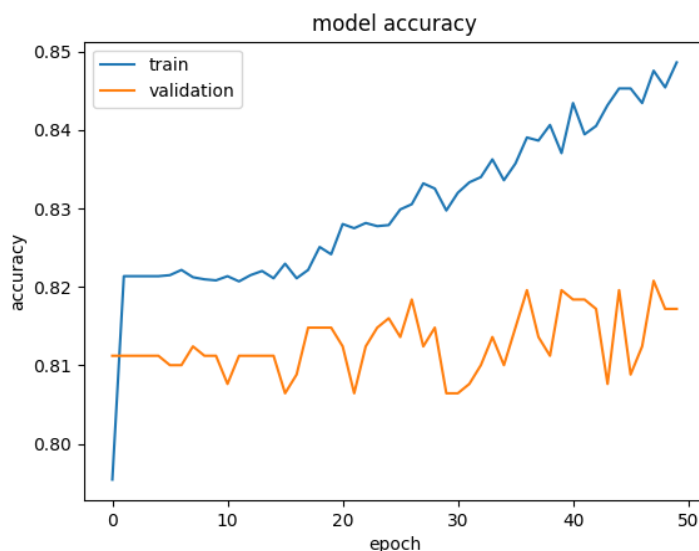
[47 1684]] falhou em 16,1%

70 neutras classificadas como neutras (3,34%)

290 neutras classificadas como não neutras (13,87%)

47 não neutras classificadas como neutras (2,27%)

1684 não neutras classificadas como não neutras (80,56%)



Accuracy do train set 85,51% é mais elevada que do validation set 81,72%

É considerado demasiado elevado se for 95% vs. 70%

Overfitting?

Stage 1: Rotation Correction Set

Test loss: 38,07%

Test accuracy: 83,12%

Train loss: 31,59%

Train accuracy: 86,06%

Validation loss: 39,90%

Validation accuracy: 80,05%

	precision	recall	f1-score	support
neutral	0.52	0.25	0.34	360
not neutral	0.86	0.95	0.90	1731
accuracy			0.83	2091
macro avg	0.69	0.60	0.62	2091
weighted avg	0.80	0.83	0.81	2091

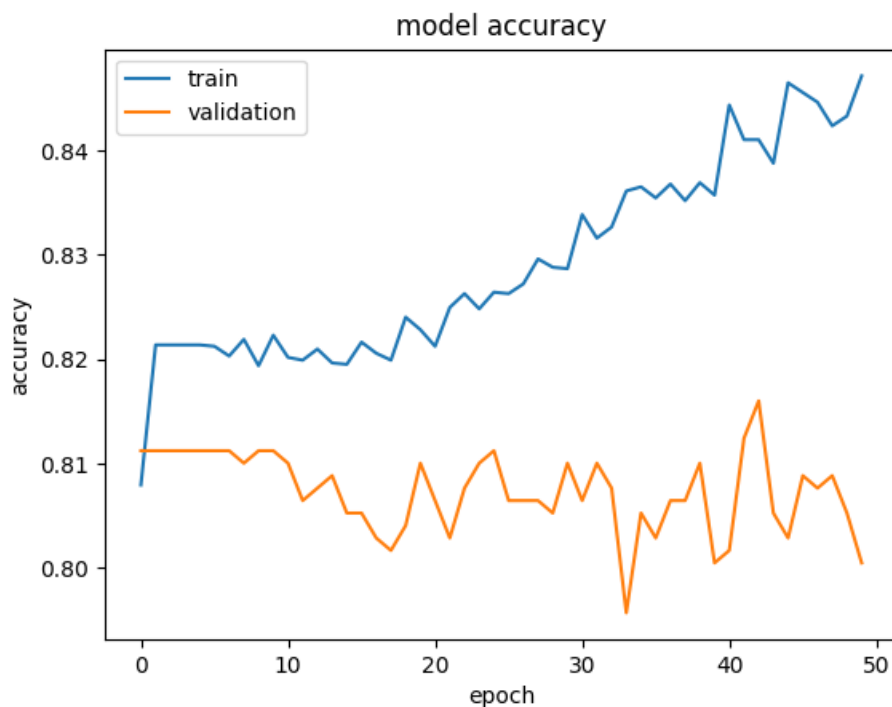
[[91 269] acertou em 83,12%
[84 1647]] falhou em 16,88%

91 neutras classificadas como neutras (4,35%)

269 neutras classificadas como não neutras (12,86%)

84 não neutras classificadas como neutras (4,02%)

1647 não neutras classificadas como não neutras (78,77%)



Accuracy do train set 86,06% é mais elevada que do validation set 80,05%

É considerado demasiado elevado se for 95% vs. 70%

Overfitting?

Stage 4: Histogram Equalization Set

Test loss: 37,60%

Test accuracy: 83,88%

Train loss: 29,81%

Train accuracy: 86,34%

Validation loss: 39,42%

Validation accuracy: 80,53%

	precision	recall	f1-score	support
neutral	0.59	0.21	0.31	360
not neutral	0.86	0.97	0.91	1731
accuracy			0.84	2091
macro avg	0.72	0.59	0.61	2091
weighted avg	0.81	0.84	0.81	2091

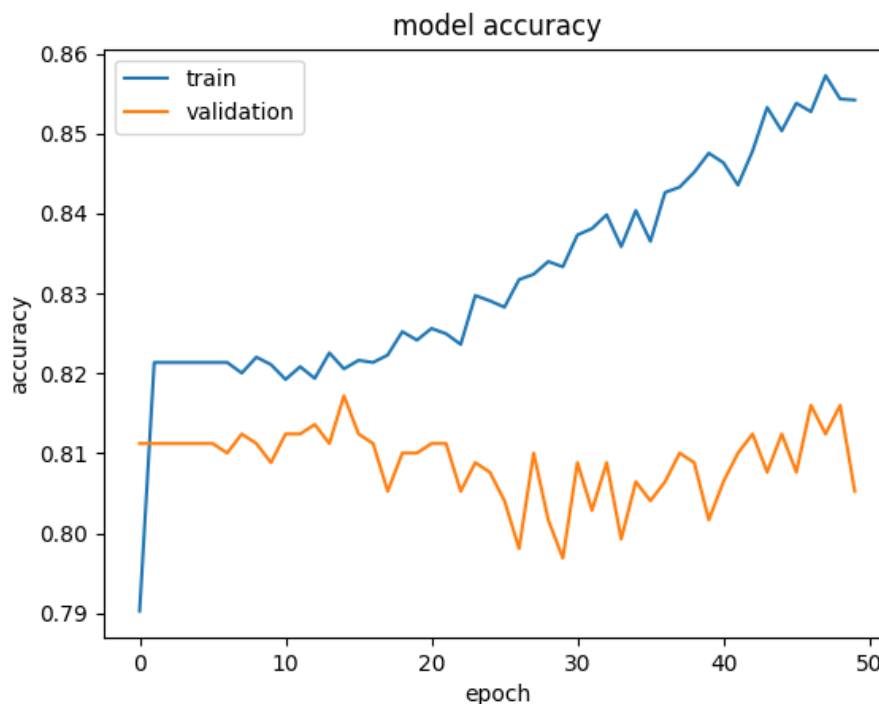
[[76 284] acertou em 83,88%
[53 1678]] falhou em 16,12%

76 neutras classificadas como neutras (3,63%)

284 neutras classificadas como não neutras (13,58%)

53 não neutras classificadas como neutras (2,53%)

1678 não neutras classificadas como não neutras (80,25%)



Accuracy do train set 86,34% é mais elevada que do validation set 80,53%

É considerado demasiado elevado se for 95% vs. 70%

Overfitting?

Stage 5: Smoothed Set

Test loss: 37,97%

Test accuracy: 82,45%

Train loss: 32,52%

Train accuracy: 85,96%

Validation loss: 40,21%

Validation accuracy: 79,45%

	precision	recall	f1-score	support
neutral	0.48	0.30	0.37	360
not neutral	0.87	0.93	0.90	1731
accuracy			0.82	2091
macro avg	0.67	0.62	0.64	2091
weighted avg	0.80	0.82	0.81	2091

[[109 251] acertou em 82,45%

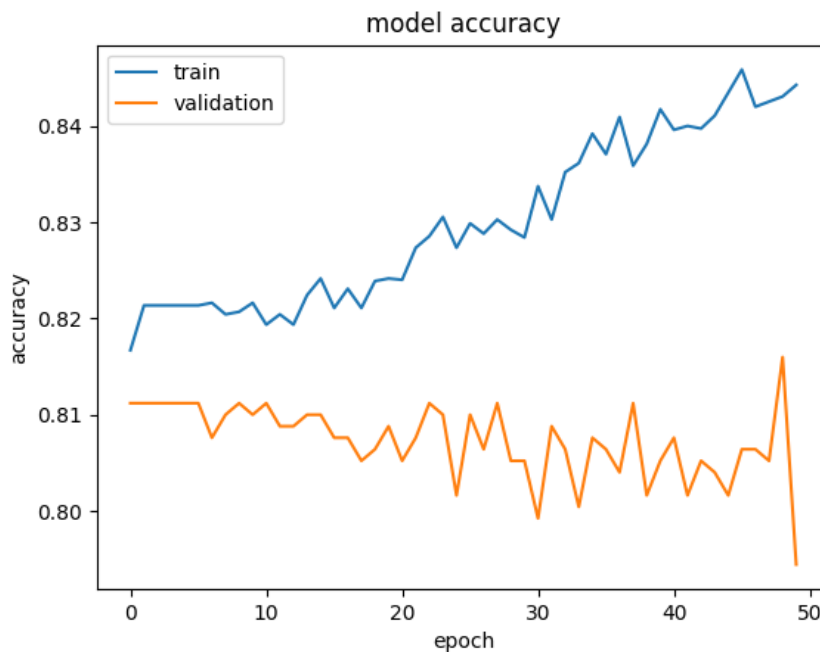
[116 1615]] falhou em 17,55%

109 neutras classificadas como neutras (5,21%)

251 neutras classificadas como não neutras (12,00%)

116 não neutras classificadas como neutras (5,55%)

1615 não neutras classificadas como não neutras (77,26%)



Accuracy do train set 85,96% é mais elevada que do validation set 79,45%

É considerado demasiado elevado se for 95% vs. 70%

Overfitting?

Model FER balanced

Impacto do Pré Processamento

Stage 0: Raw Set

Test loss: 53,88%

Test accuracy: 73,44%

Train loss: 43,53%

Train accuracy: 80,63%

Validation loss: 55,42%

Validation accuracy: 72,20%

	precision	recall	f1-score	support
neutral	0.71	0.82	0.76	374
not neutral	0.77	0.65	0.71	364
accuracy			0.73	738
macro avg	0.74	0.73	0.73	738
weighted avg	0.74	0.73	0.73	738

[[305 69] acertou em 73,44%

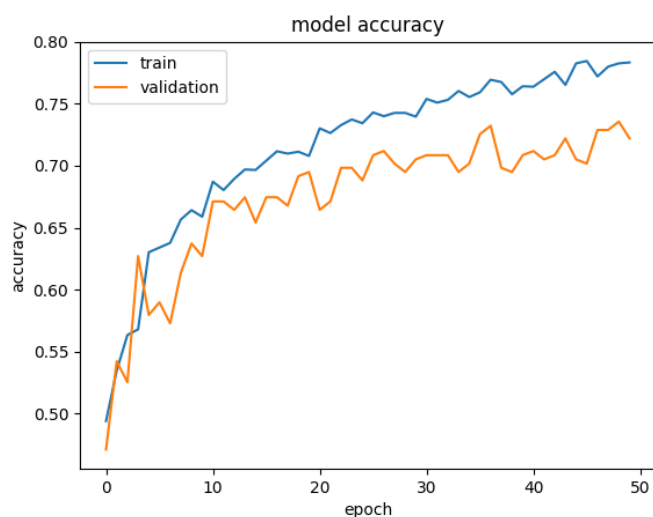
[127 237]] falhou em 26,56%

305 neutras classificadas como neutras (41,33%)

69 neutras classificadas como não neutras (9,35%)

127 não neutras classificadas como neutras (17,21%)

237 não neutras classificadas como não neutras (32,11%)



Accuracy do train set 80,63% é mais elevada que do validation set 72,20%

É considerado demasiado elevado se for 95% vs. 70%

Overfitting?

Stage 1: Rotation Correction Set

Test loss: 53,67%

Test accuracy: 73,44%

Train loss: 44,67%

Train accuracy: 80,66%

Validation loss: 50,92%

Validation accuracy: 74,24%

	precision	recall	f1-score	support
neutral	0.72	0.78	0.75	374
not neutral	0.75	0.69	0.72	364
accuracy			0.73	738
macro avg	0.74	0.73	0.73	738
weighted avg	0.74	0.73	0.73	738

[[290 84] acertou em 73,44%

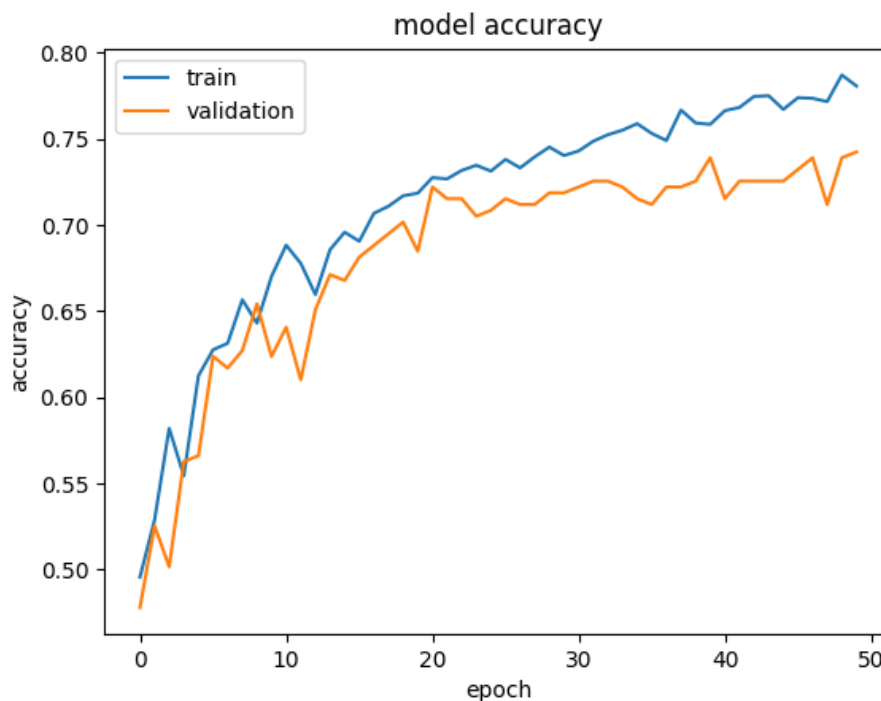
[112 252]] falhou em 26,56%

290 neutras classificadas como neutras (39,30%)

84 neutras classificadas como não neutras (11,38%)

112 não neutras classificadas como neutras (15,18%)

252 não neutras classificadas como não neutras (34,15%)



Accuracy do train set 80,66% é mais elevada que do validation set 74,24%
É considerado demasiado elevado se for 95% vs. 70%

Stage 4: Histogram Equalization Set

Test loss: 52,19%

Test accuracy: 74,25%

Train loss: 42,31%

Train accuracy: 81,45%

Validation loss: 50,96%

Validation accuracy: 73,56%

	precision	recall	f1-score	support
neutral	0.72	0.81	0.76	374
not neutral	0.78	0.67	0.72	364
accuracy			0.74	738
macro avg	0.75	0.74	0.74	738
weighted avg	0.75	0.74	0.74	738

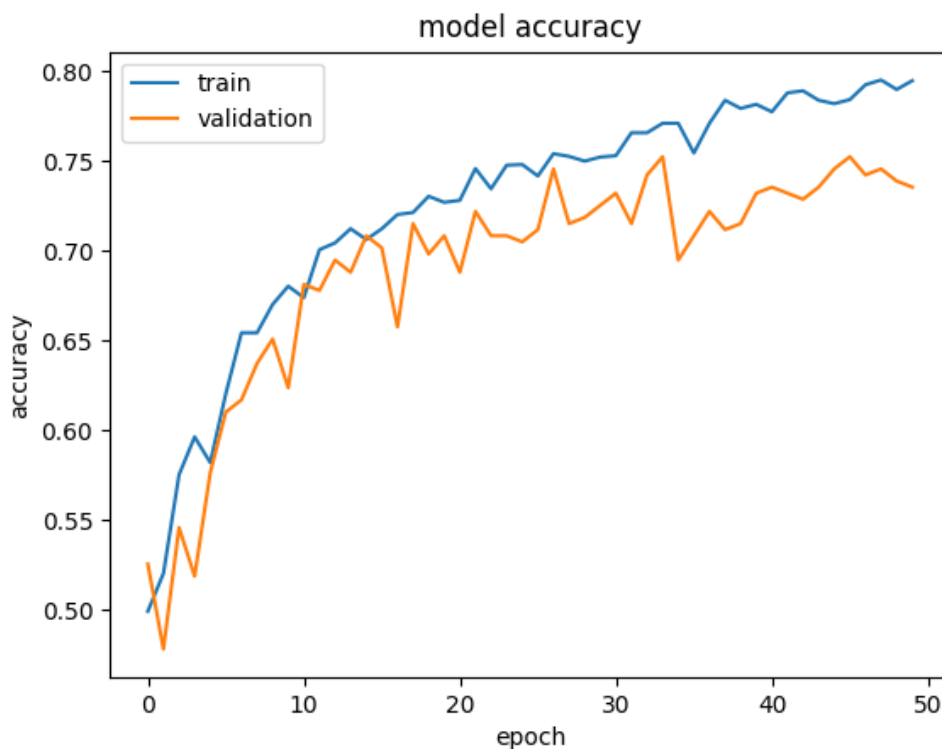
[[304 70] acertou em 74,25%
[120 244]] falhou em 25,75%

304 neutras classificadas como neutras (41,19%)

70 neutras classificadas como não neutras (9,49%)

120 não neutras classificadas como neutras (16,26%)

244 não neutras classificadas como não neutras (33,06%)



Accuracy do train set 81,45% é mais elevada que do validation set 73,56%
É considerado demasiado elevado se for 95% vs. 70%

Stage 5: Smoothed Set

Test loss: 53,74%

Test accuracy: 74,66%

Train loss: 47,87%

Train accuracy: 77,46%

Validation loss: 51,17%

Validation accuracy: 74,57%

	precision	recall	f1-score	support
neutral	0.72	0.81	0.76	374
not neutral	0.78	0.68	0.73	364
accuracy			0.75	738
macro avg	0.75	0.75	0.75	738
weighted avg	0.75	0.75	0.75	738

[[304 70] acertou em 74,66%

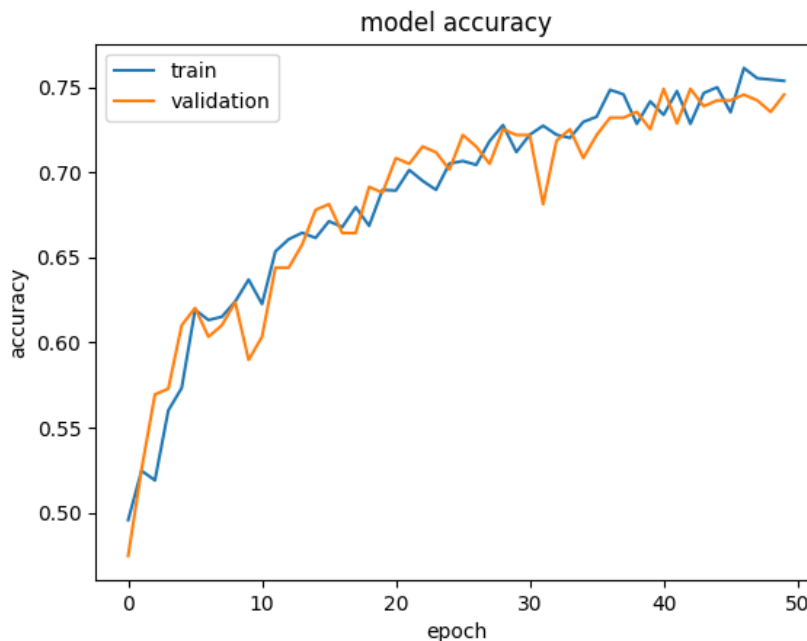
[117 247]] falhou em 25,34%

304 neutras classificadas como neutras (41,19%)

70 neutras classificadas como não neutras (9,49%)

117 não neutras classificadas como neutras (15,85%)

247 não neutras classificadas como não neutras (33,47%)



Accuracy do train set 77,46% é muito próximo do validation set 74,57%

is not overfitting, but it is not performing optimally either

could be that the model is not being trained for enough iterations

validation set is not large enough

model's capacity is not large enough to fully learn the underlying patterns in the data

Model CK+

Impacto do Pré-Processamento

Stage 0: Raw Set

Test loss: 69,72%

Test accuracy: 82,61%

Train loss: 2,73%

Train accuracy: 99,87%

Validation loss: 96,66%

Validation accuracy: 77,11%

	precision	recall	f1-score	support
neutral	0.87	0.84	0.86	57
not neutral	0.76	0.80	0.78	35
accuracy			0.83	92
macro avg	0.81	0.82	0.82	92
weighted avg	0.83	0.83	0.83	92

[[48 9] acertou em 82,61%

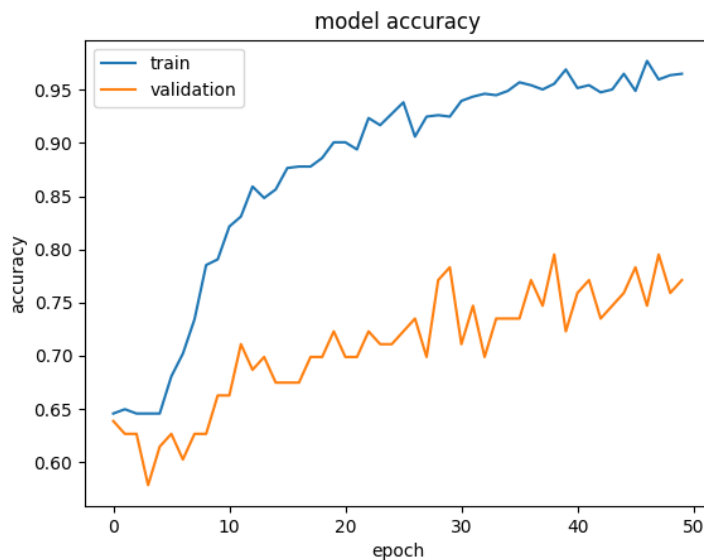
[7 28]] falhou em 17,39%

48 neutras classificadas como neutras (52,17%)

9 neutras classificadas como não neutras (9,78%)

7 não neutras classificadas como neutras (7,61%)

28 não neutras classificadas como não neutras (30,43%)



Accuracy do train set 99,87% é muito mais elevada que do validation set 77,11%

Overfitting

Stage 1: Rotation Correction Set

Test loss: 84,04%

Test accuracy: 70,65%

Train loss: 2,71%

Train accuracy: 99,46%

Validation loss: 96,47%

Validation accuracy: 68,67%

	precision	recall	f1-score	support
neutral	0.74	0.81	0.77	57
not neutral	0.63	0.54	0.58	35
accuracy			0.71	92
macro avg	0.69	0.67	0.68	92
weighted avg	0.70	0.71	0.70	92

[[46 11] acertou em 70,65%

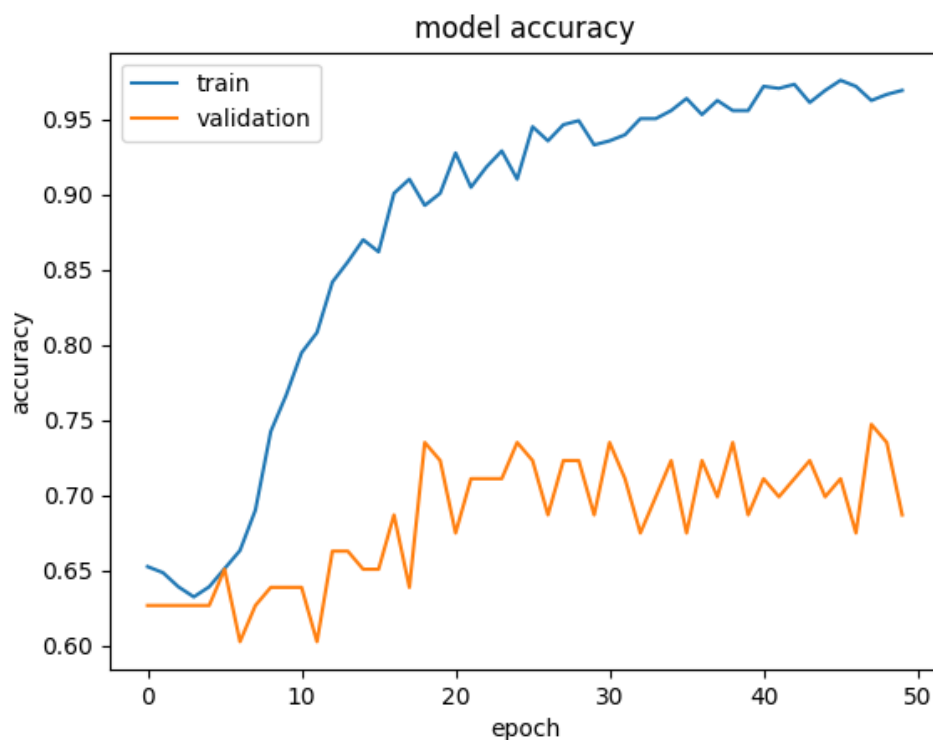
[16 19]] falhou em 29,35%

46 neutras classificadas como neutras (50,00%)

11 neutras classificadas como não neutras (11,96%)

16 não neutras classificadas como neutras (17,39%)

19 não neutras classificadas como não neutras (20,65%)



Accuracy do train set 99,46% é muito mais elevada que do validation set 68,67%
Overfitting

Stage 2: Cropped Set

Test loss: 46,65%

Test accuracy: 88,04%

Train loss: 3,72%

Train accuracy: 99,06%

Validation loss: 2,11%

Validation accuracy: 92,11%

	precision	recall	f1-score	support
neutral	0.90	0.91	0.90	57
not neutral	0.85	0.83	0.84	35
accuracy			0.88	92
macro avg	0.87	0.87	0.87	92
weighted avg	0.88	0.88	0.88	92

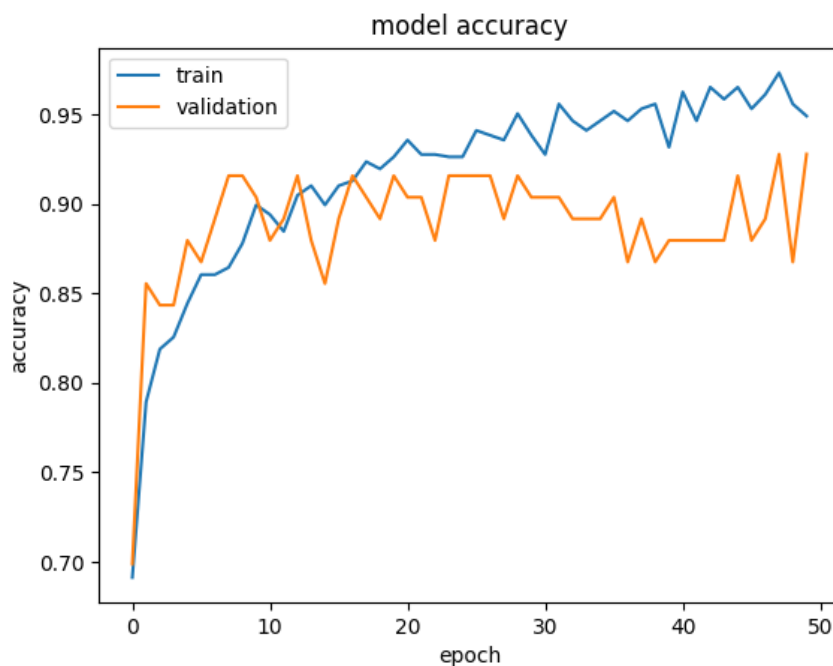
[[52 5] acertou em 88,04%
[6 29]] falhou em 11,96%

52 neutras classificadas como neutras (56,52%)

5 neutras classificadas como não neutras (5,43%)

6 não neutras classificadas como neutras (6,52%)

29 não neutras classificadas como não neutras (31,52%)



Accuracy do train set 99,06% é mais elevada que do validation set 92,11%

is not overfitting, but it is not performing optimally either

could be that the model is not being trained for enough iterations

validation set is not large enough

model's capacity is not large enough to fully learn the underlying patterns in the data

Stage 4: Histogram Equalization Set

Test loss: 35,31%

Test accuracy: 94,56%

Train loss: 0,26%

Train accuracy: 100%

Validation loss: 25,28%

Validation accuracy: 90,36%

	precision	recall	f1-score	support
neutral	0.93	0.98	0.96	57
not neutral	0.97	0.89	0.93	35
accuracy			0.95	92
macro avg	0.95	0.93	0.94	92
weighted avg	0.95	0.95	0.95	92

[[56 1] acertou em 94,56%

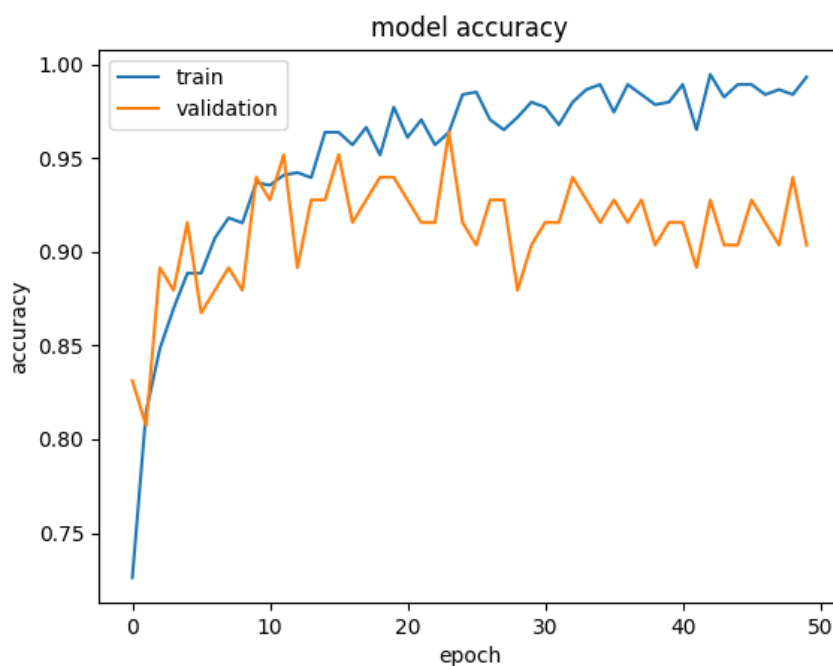
[4 31]] falhou em 5,44%

56 neutras classificadas como neutras (60,87%)

1 neutras classificadas como não neutras (1,09%)

4 não neutras classificadas como neutras (4,35%)

31 não neutras classificadas como não neutras (33,70%)



Accuracy do train set 100% é mais elevada que do validation set 90,36%

is not overfitting, but it is not performing optimally either

could be that the model is not being trained for enough iterations

validation set is not large enough

model's capacity is not large enough to fully learn the underlying patterns in the data

Stage 5: Smoothed Set

Test loss: 38,90%

Test accuracy: 92,39%

Train loss: 1,31%

Train accuracy: 99,73%

Validation loss: 18,36%

Validation accuracy: 93,98%

	precision	recall	f1-score	support
neutral	0.90	0.98	0.94	57
not neutral	0.97	0.83	0.89	35
accuracy			0.92	92
macro avg	0.93	0.91	0.92	92
weighted avg	0.93	0.92	0.92	92

[[56 1] acertou em 92,39%

[6 29]] falhou em 7,61%

56 neutras classificadas como neutras (60,87%)

1 neutras classificadas como não neutras (1,09%)

6 não neutras classificadas como neutras (6,52%)

29 não neutras classificadas como não neutras (31,52%)



Accuracy do train set 99,73% é mais elevada que do validation set 93,98%

is not overfitting, but it is not performing optimally either

could be that the model is not being trained for enough iterations

validation set is not large enough

model's capacity is not large enough to fully learn the underlying patterns in the data

Model CK+ balanceado

Impacto do Pré-Processamento

Stage 0: Raw Set

Test loss: 76,58%

Test accuracy: 43,94%

Train loss: 63,33%

Train accuracy: 64,84%

Validation loss: 68,26%

Validation accuracy: 54,24%

	precision	recall	f1-score	support
neutral	0.46	0.56	0.51	34
not neutral	0.40	0.31	0.35	32
accuracy			0.44	66
macro avg	0.43	0.44	0.43	66
weighted avg	0.43	0.44	0.43	66

[[19 15] acertou em 43,94%

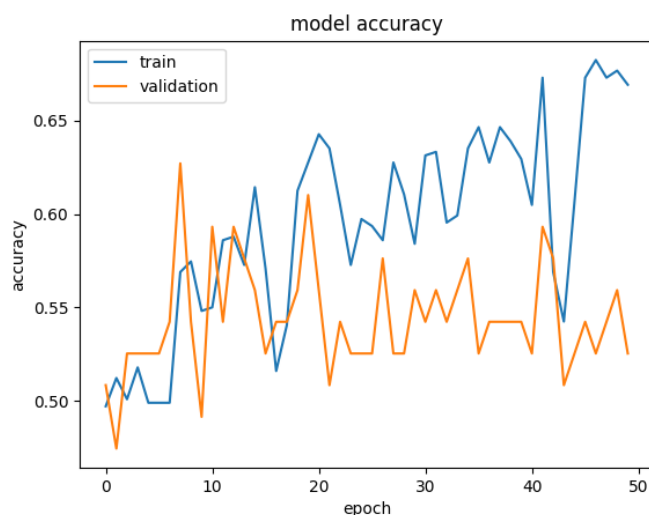
[22 10]] falhou em 56,06%

19 neutras classificadas como neutras (28,78%)

15 neutras classificadas como não neutras (22,73%)

22 não neutras classificadas como neutras (33,33%)

10 não neutras classificadas como não neutras (15,15%)



Accuracy do train set 64,84% é mais elevada que do validation set 54,24%

is not overfitting, but it is not performing optimally either

could be that the model is not being trained for enough iterations

validation set is not large enough

model's capacity is not large enough to fully learn the underlying patterns in the data

Stage 1: Rotation Correction Set

Test loss: 81,61%

Test accuracy: 48,48%

Train loss: 66,36%

Train accuracy: 57,47%

Validation loss: 74,74%

Validation accuracy: 50,85%

	precision	recall	f1-score	support
neutral	0.50	0.82	0.62	34
not neutral	0.40	0.12	0.19	32
accuracy			0.48	66
macro avg	0.45	0.47	0.41	66
weighted avg	0.45	0.48	0.41	66

[[28 6] acertou em 48,48%

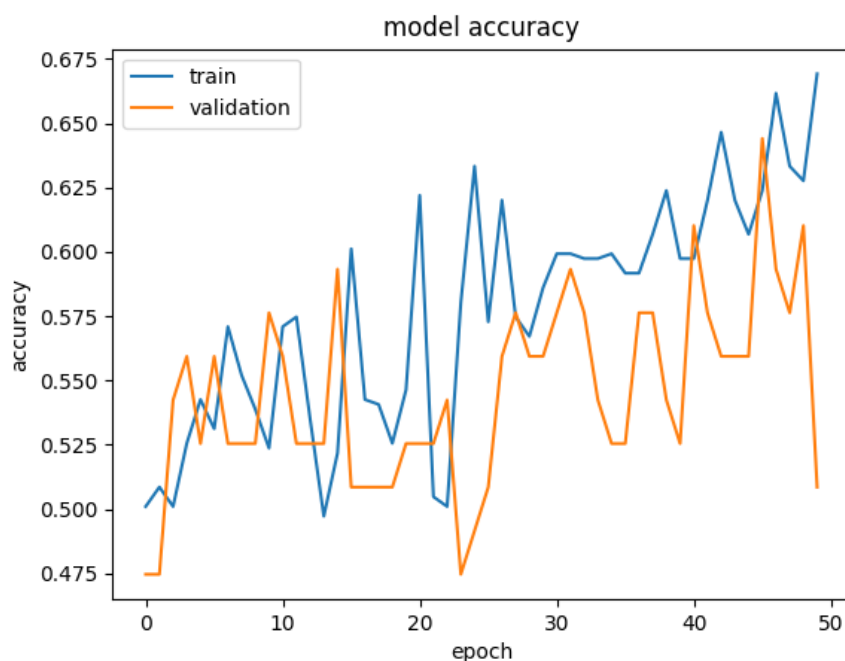
[28 4]] falhou em 51,52%

28 neutras classificadas como neutras (42,42%)

6 neutras classificadas como não neutras (9,09%)

28 não neutras classificadas como neutras (42,42%)

4 não neutras classificadas como não neutras (6,06%)



Accuracy do train set 57,47% é mais elevada que do validation set 50,85%

is not overfitting, but it is not performing optimally either

could be that the model is not being trained for enough iterations

validation set is not large enough

model's capacity is not large enough to fully learn the underlying patterns in the data

Stage 2: Cropped Set

Test loss: 35,09%

Test accuracy: 86,36%

Train loss: 34,67%

Train accuracy: 84,69%

Validation loss: 37,44%

Validation accuracy: 81,36%

	precision	recall	f1-score	support
neutral	0.86	0.88	0.87	34
not neutral	0.87	0.84	0.86	32
accuracy			0.86	66
macro avg	0.86	0.86	0.86	66
weighted avg	0.86	0.86	0.86	66

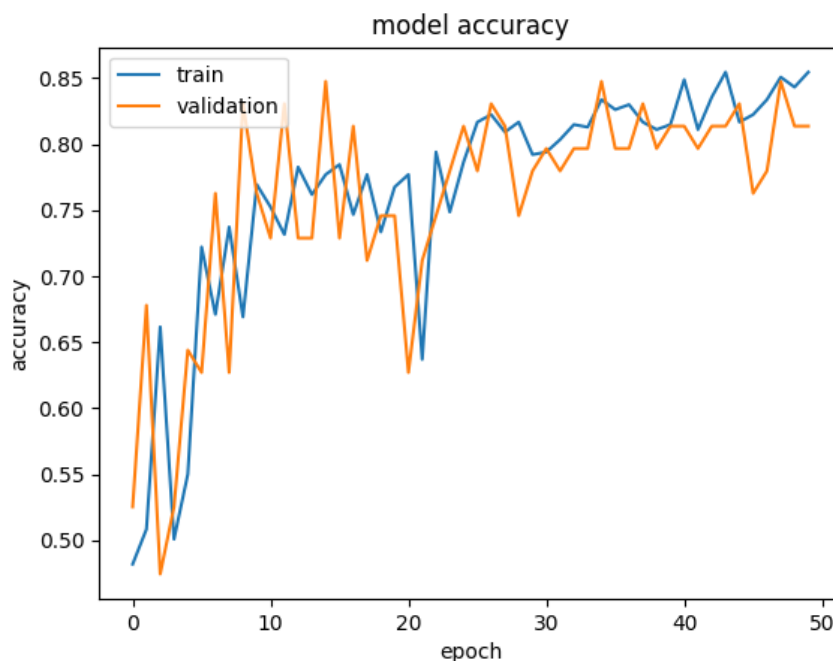
[[30 4] acertou em 86,36%
[5 27]] falhou em 13,64%

30 neutras classificadas como neutras (45,45%)

4 neutras classificadas como não neutras (6,06%)

5 não neutras classificadas como neutras (7,58%)

27 não neutras classificadas como não neutras (40,91%)



Accuracy do train set 84,69% é mais elevada que do validation set 81,36%

is not overfitting, but it is not performing optimally either

could be that the model is not being trained for enough iterations

validation set is not large enough

model's capacity is not large enough to fully learn the underlying patterns in the data

Stage 4: Histogram Equalization Set

Test loss: 32,24%

Test accuracy: 84,85%

Train loss: 32,90%

Train accuracy: 84,69%

Validation loss: 44,66%

Validation accuracy: 79,66%

	precision	recall	f1-score	support
neutral	0.79	0.97	0.87	34
not neutral	0.96	0.72	0.82	32
accuracy			0.85	66
macro avg	0.87	0.84	0.84	66
weighted avg	0.87	0.85	0.85	66

[[33 1] acertou em 84,85%

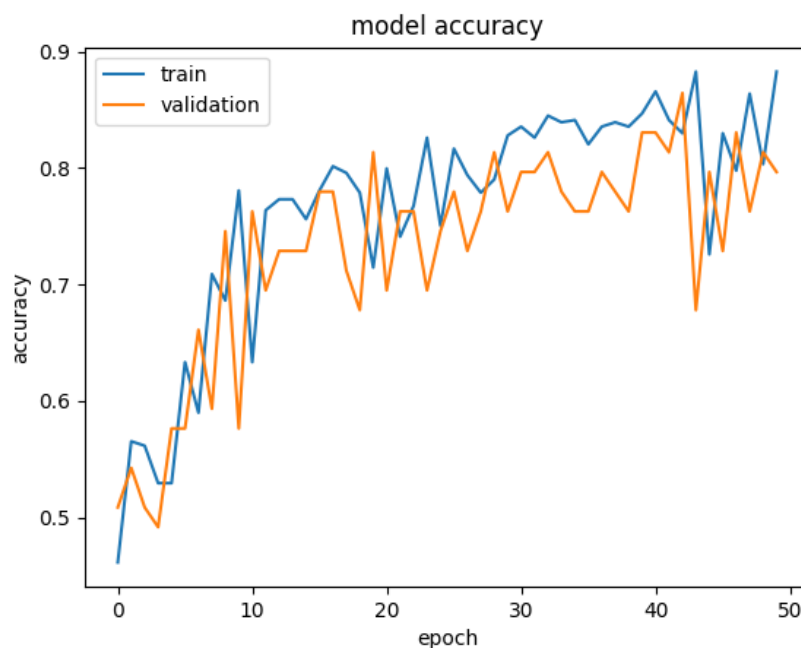
[9 23]] falhou em 15,15%

33 neutras classificadas como neutras (50,00%)

1 neutras classificadas como não neutras (1,52%)

9 não neutras classificadas como neutras (13,64%)

23 não neutras classificadas como não neutras (34,85%)



Accuracy do train set 84,69% é mais elevada que do validation set 79,66%

is not overfitting, but it is not performing optimally either

could be that the model is not being trained for enough iterations

validation set is not large enough

model's capacity is not large enough to fully learn the underlying patterns in the data

Stage 5: Smoothed Set

Test loss: 41,07%

Test accuracy: 80,30%

Train loss: 40,70%

Train accuracy: 80,15%

Validation loss: 54,57%

Validation accuracy: 72,88%

	precision	recall	f1-score	support
neutral	0.73	0.97	0.84	34
not neutral	0.95	0.62	0.75	32
accuracy			0.80	66
macro avg	0.84	0.80	0.80	66
weighted avg	0.84	0.80	0.80	66

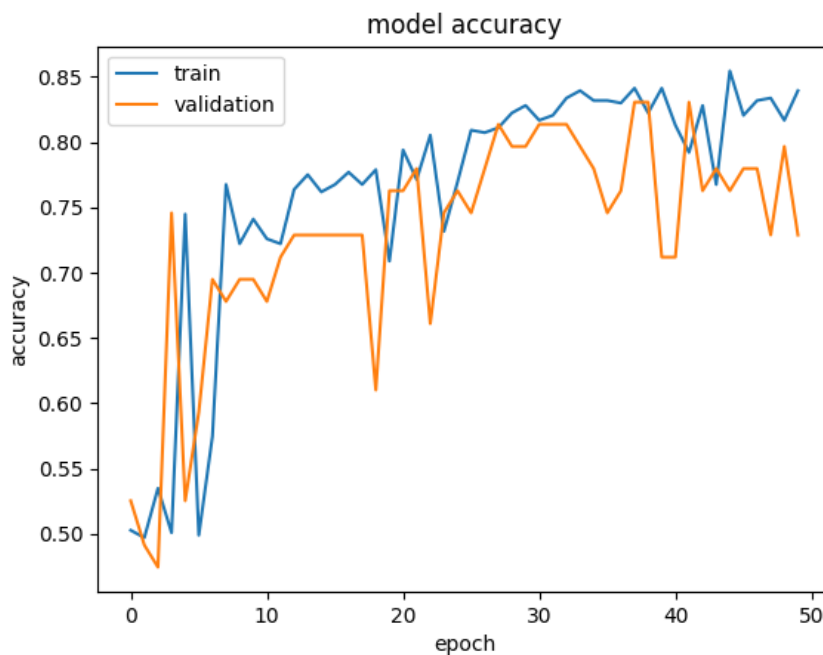
[[33 1] acertou em 80,30%
[12 20]] falhou em 19,70%

33 neutras classificadas como neutras (50,00%)

1 neutras classificadas como não neutras (1,52%)

12 não neutras classificadas como neutras (18,18%)

20 não neutras classificadas como não neutras (30,30%)



Accuracy do train set 80,15% é mais elevada que do validation set 72,88%

is not overfitting, but it is not performing optimally either

could be that the model is not being trained for enough iterations

validation set is not large enough

model's capacity is not large enough to fully learn the underlying patterns in the data

Model CK+ data augmentation

Impacto do Pré-Processamento

Stage 0: Raw Set

Test loss: 45,78%

Test accuracy: 82,95%

Train loss: 37,97%

Train accuracy: 87,91%

Validation loss: 48,35%

Validation accuracy: 79,49%

	precision	recall	f1-score	support
neutral	0.81	0.90	0.85	118
not neutral	0.86	0.75	0.80	99
accuracy			0.83	217
macro avg	0.83	0.82	0.83	217
weighted avg	0.83	0.83	0.83	217

[[106 12] acertou em 82,95%

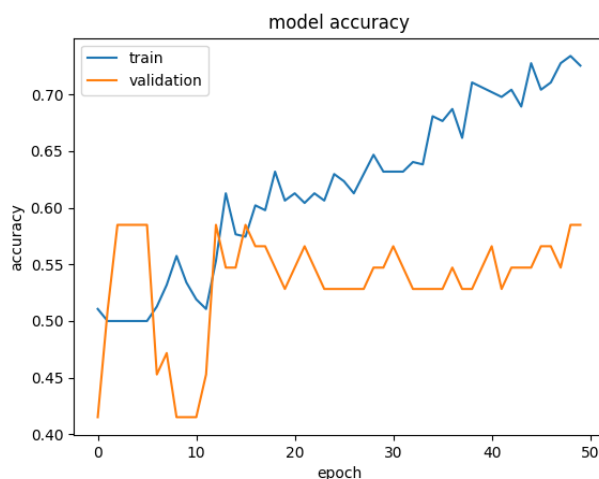
[25 74]] falhou em 17,05%

106 neutras classificadas como neutras (48,85%)

12 neutras classificadas como não neutras (5,53%)

25 não neutras classificadas como neutras (11,52%)

74 não neutras classificadas como não neutras (34,10%)



Accuracy do train set 87,91% é mais elevada que do validation set 79,49%

is not overfitting, but it is not performing optimally either

could be that the model is not being trained for enough iterations

validation set is not large enough

model's capacity is not large enough to fully learn the underlying patterns in the data

Stage 1: Rotation Correction Set

Test loss: 46,35%

Test accuracy: 76,95%

Train loss: 37,01%

Train accuracy: 86,71%

Validation loss: 48,07%

Validation accuracy: 77,44%

	precision	recall	f1-score	support
neutral	0.81	0.75	0.78	118
not neutral	0.73	0.79	0.76	99
accuracy			0.77	217
macro avg	0.77	0.77	0.77	217
weighted avg	0.77	0.77	0.77	217

[[89 29] acertou em 76,95%

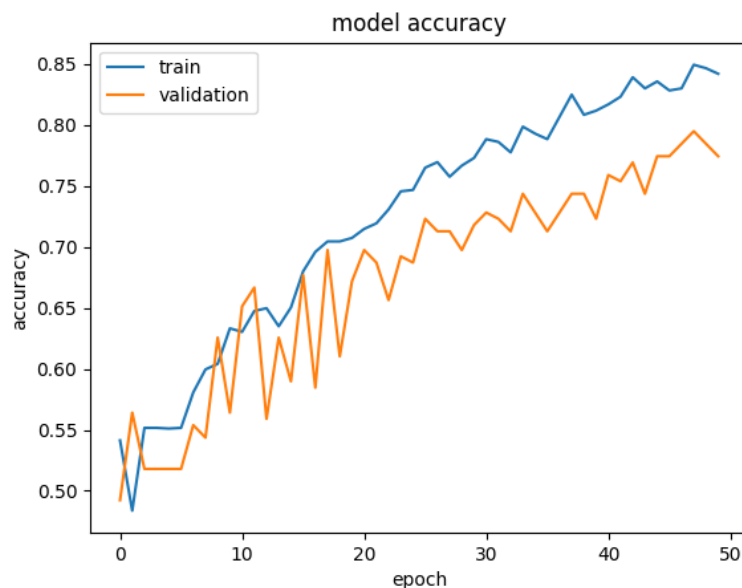
[21 78]] falhou em 23,05%

89 neutras classificadas como neutras (41,01%)

29 neutras classificadas como não neutras (13,36%)

21 não neutras classificadas como neutras (9,68%)

78 não neutras classificadas como não neutras (35,95%)



Accuracy do train set 86,71% é mais elevada que do validation set 77,44%

is not overfitting, but it is not performing optimally either

could be that the model is not being trained for enough iterations

validation set is not large enough

model's capacity is not large enough to fully learn the underlying patterns in the data

Stage 2: Cropped Set

Test loss: 23,33%

Test accuracy: 89,40%

Train loss: 21,63%

Train accuracy: 91,56%

Validation loss: 22,25%

Validation accuracy: 91,28%

	precision	recall	f1-score	support
neutral	0.88	0.93	0.91	118
not neutral	0.91	0.85	0.88	99
accuracy			0.89	217
macro avg	0.90	0.89	0.89	217
weighted avg	0.90	0.89	0.89	217

[[110 8] acertou em 89,40%

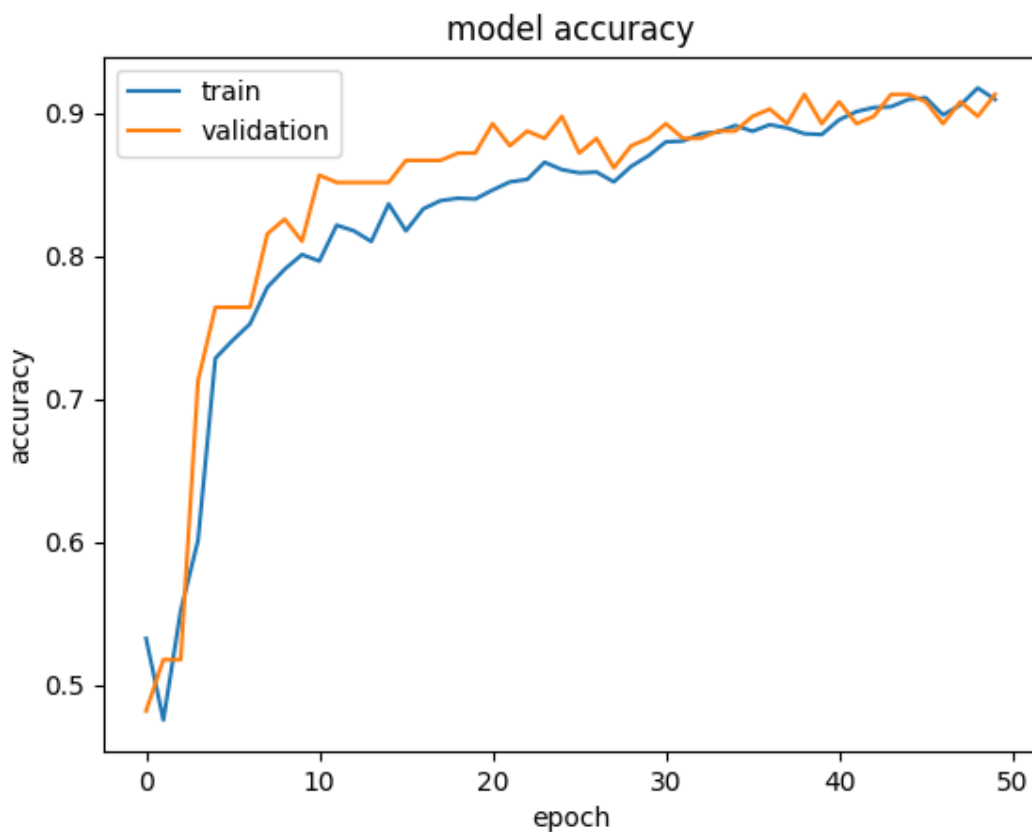
[15 84]] falhou em 10,60%

110 neutras classificadas como neutras (50,69%)

8 neutras classificadas como não neutras (3,69%)

15 não neutras classificadas como neutras (6,91%)

84 não neutras classificadas como não neutras (38,71%)



Overfitting

Stage 4: Histogram Equalization Set

Test loss: 17,51%

Test accuracy: 93,09%

Train loss: 13,44%

Train accuracy: 95,89%

Validation loss: 18,83%

Validation accuracy: 93,33%

	precision	recall	f1-score	support
neutral	0.92	0.96	0.94	118
not neutral	0.95	0.90	0.92	99
accuracy			0.93	217
macro avg	0.93	0.93	0.93	217
weighted avg	0.93	0.93	0.93	217

[[113 5] acertou em 93,09%

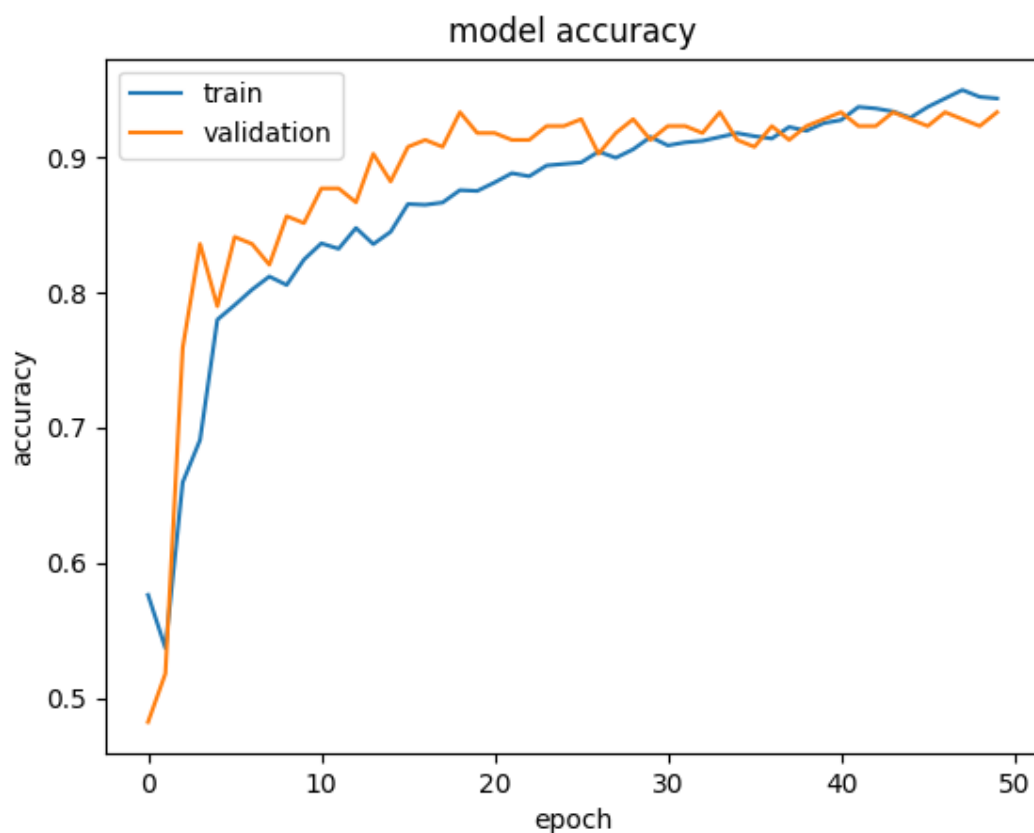
[10 89]] falhou em 6,91%

113 neutras classificadas como neutras (52,07%)

5 neutras classificadas como não neutras (2,30%)

10 não neutras classificadas como neutras (4,61%)

89 não neutras classificadas como não neutras (41,01%)



Overfitting

Stage 5: Smoothed Set

Test loss: 20,86%

Test accuracy: 92,17%

Train loss: 16,76%

Train accuracy: 93,78%

Validation loss: 20,30%

Validation accuracy: 93,33%

	precision	recall	f1-score	support
neutral	0.89	0.97	0.93	118
not neutral	0.97	0.86	0.91	99
accuracy			0.92	217
macro avg	0.93	0.92	0.92	217
weighted avg	0.93	0.92	0.92	217

[[115 3] acertou em 92,17%

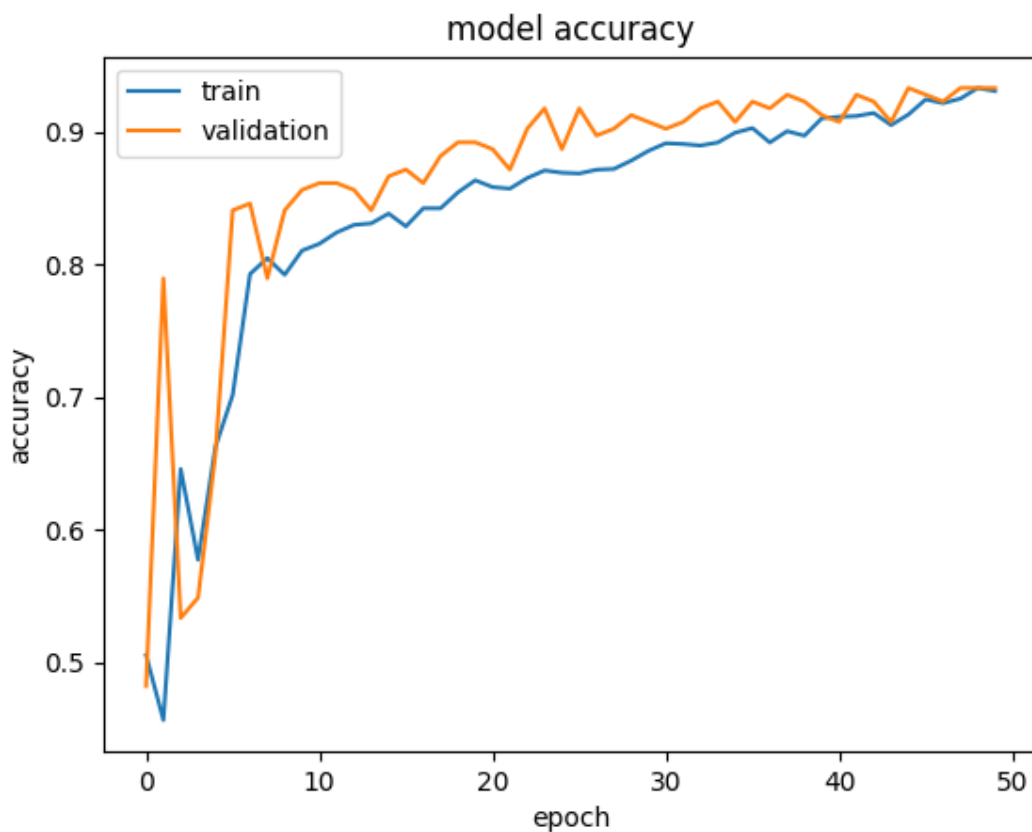
[14 85]] falhou em 7,83%

115 neutras classificadas como neutras (53,00%)

3 neutras classificadas como não neutras (1,38%)

14 não neutras classificadas como neutras (6,45%)

85 não neutras classificadas como não neutras (39,17%)



Overfitting

Stage 0: Raw Set FOR OTHER DATA AUGMENTATION (instead of changing contrast do Random horizontal flipping: flip the image horizontally with a probability of 50%)

Number of images flipped: 380

Test loss: 87,17%

Test accuracy: 58,70%

Train loss: 48,68%

Train accuracy: 76,24%

Validation loss: 84,07%

Validation accuracy: 61,46%

	precision	recall	f1-score	support
neutral	0.62	0.88	0.72	57
not neutral	0.36	0.11	0.17	35
accuracy			0.59	92
macro avg	0.49	0.50	0.45	92
weighted avg	0.52	0.59	0.52	92

[[50 7]

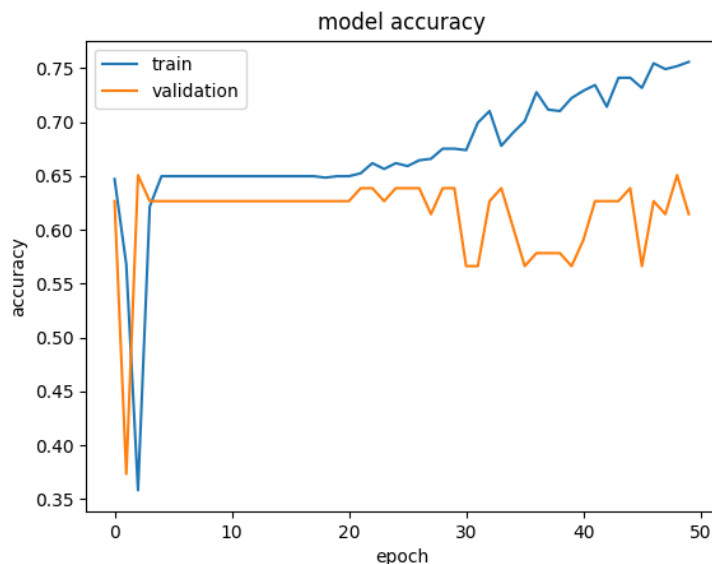
[31 4]]

50 neutras classificadas como neutras

7 neutras classificadas como não neutras

31 não neutras classificadas como neutras

4 não neutras classificadas como não neutras



Piores resultados mais não neutras a serem classificadas como neutras

The **accuracy of the training set** is a measure of how well the model is able to fit the **training data**. A high accuracy on the training set indicates that the model is able to fit the training data well and has low bias. However, it does not necessarily mean that the model will perform well on new, unseen data.

The **accuracy of the test set** is a measure of how well the model is able to generalize to **new, unseen data**. A high accuracy on the test set indicates that the model is able to

generalize well and has low variance. It means that the model is likely to perform well on unseen data too.

The **accuracy of the validation set**, also known as a hold-out set, is a measure of how well the model is able to generalize to new, unseen data. It is used to estimate the performance of a model when it is used to make new predictions. Like the test set, a high accuracy on the validation set indicates that the model is likely to perform well on unseen data.

Generally, we expect the **accuracy of the training set** to be **higher** than the **accuracy of the validation set**, as the model is overfitting to the training data. In the best cases, accuracy of the validation set would be close to the accuracy of the training set, which means that the model is able to generalize well to new data.

In addition, **overfitting** occurs when the **accuracy of the training set** is **much higher** than the accuracy of the validation set. This means the model is very good at fitting the training data but not as good at generalizing to new data. In this case, we can try different techniques such as regularization or adding more data to the training set to improve the model's performance.

In machine learning, loss is a function that measures the difference between the predicted output and the actual output for a given set of inputs. Loss is commonly used to train machine learning models, and the goal is to minimize it.

A common way to visualize loss is through a graph, such as a line plot or a scatter plot, of the loss on the train, test, and validation sets as the model is trained and the number of iterations increase. This can help us identify trends in the loss over time and make decisions about when to stop training or change the model architecture.

In general, the loss of the training set should decrease as the number of iterations increases. Ideally, the loss of the validation set should decrease as well, but there's a chance that it will increase after a number of iterations, this is because it may be overfitting to the training set. If the **loss of the validation set increases**, this may be an indication that the model is overfitting to the training set.

The loss of the test set should also be close to the loss of the validation set and not significantly larger, this is because the test set is unseen data and it should be representative of the model generalization.

It is important to note that the accuracy and loss are two different measures, the accuracy is the percentage of correct predictions and the loss is a measure of how far the predictions are from the true values. Sometimes the accuracy of the model may be high while the loss is high or vice versa, so they should be evaluated together to make better decisions.