United International Journal for Research & Technology
Volume 06, Issue 03, 2024 | Open Access | ISSN: 2582-6832

An Enhancement of Gated Recurrent Unit (GRU) for Speech Emotion Recognition in the Implementation of Voice-Based Danger Recognition System

Jayson C. Sepeda[1], Louise Dawn F. Santos[2], and Prof. Leisyl Mahusay[3]
[1,2]Student, College of Information Systems and Technology Management, Pamantasan ng Lungsod ng Maynila (University of the City of Manila), Manila Philippines
[3]Faculty Adviser, College of Information Systems and Technology Management, Pamantasan ng Lungsod ng Maynila (University of the City of Manila), Manila

Abstract**—** As technology grows, fields like Speech Emotion Recognition (SER) are enriched. SER, which uses speech signals to recognize human emotions, is used in various services. Usage of Gated Recurrent Units (GRU) is prominent in SER. However, GRU faces the problem of overfitting where the model fails to properly generalize and fits too closely to the training data, causing poor performance on unseen data. The aim of this study is to simulate overfitting, and to solve it effectively. Different optimizations were tested to see what solution would best overcome the overfitting problem and to improve the test accuracy, the measurement of how the model performs on unseen data. Out of all the solutions, the combination of Dropout (20%), Batch normalization, and Xavier/Glorot Initialization produced the best improvements. The base model was able to produce only 63.21% accuracy on test data, with overfitting arising. Meanwhile, the optimized model was able to mitigate overfitting and was able to increase the test accuracy to 67.34%. This concludes that the enhanced model is less prone to overfitting which results in higher test accuracy. Furthermore, the enhanced model is implemented in a danger recognition system that aims to strengthen safety, especially for the speech impaired.

Keywords**—** Speech Emotion Recognition (SER), Gated Recurrent Units (GRU), Overfitting, Danger Recognition System, GRU Optimization.

I. INTRODUCTION

As the global society progresses, the usage of Deep Learning has become more and more relevant. It is now a complete foundation for technological pieces that have been able to change the trajectory of the world. Deep learning is a type of artificial intelligence (AI) technology that teaches machines to interpret information similarly to the human brain. Following that, a type of deep learning architecture that is specifically designed for sequential data processing tasks is the Gated Recurrent Unit (GRU) algorithm. GRU is a type of recurrent neural network (RNN), and was created as a simpler alternative to Long Short-Term Memory (LSTM) networks in 2014 by Kyunghyun Cho et al. GRU's process is to use gating mechanisms to selectively update the hidden state of the network at each time step. Information entering and leaving the network is managed by the gating mechanisms, which are the reset gate and update gate. The functionalities and applications of the GRU algorithm can be applied in various technological fields, such as Natural Language Processing (NLP), Time Series Prediction, Image Analysis, and Speech Recognition.

This study aims to enhance the current Gated Recurrent Unit (GRU) algorithm for Speech Emotion Recognition (SER) in the implementation of a Voice-Based Danger Recognition System. The goal of using Speech Emotion Recognition is to deduce the emotional state of the speaker—happiness, anger, sadness, or fear—from their speech patterns, which include prosody, pitch, and rhythm. Incorporating the power of Speech Emotion Recognition (SER) with the capabilities of the Gated Recurrent Unit (GRU) algorithm is the purpose of this study, which will result in the creation of a system that recognizes if a person is in danger through their voice and tone. With that purpose, the study will mainly benefit PWDs (particularly persons with speech impediment) and victims of danger. Reportedly, the Philippines' 911 Emergency hotline received 18.4 million calls in 2019, which is a major jump from 2016's 1.46 million calls. This shows that the number of crimes and emergencies reported increases continuously. Unfortunately, the dangers and threats experienced by Persons With Disabilities in the Philippines are worse, especially those who have disabilities that hinders them to articulate their concern in the form of speech. Particularly, those who are put in a threatening position that makes it impossible for them to speak as it puts their

United International Journal for Research & Technology
Volume 06, Issue 03, 2024 | Open Access | ISSN: 2582-6832

safety and lives on the line. According to the Philippines Statistics Authority, Persons with functional difficulty are posted at 8.7%, which is nearly nine in every 100 persons. Additionally, data from Gitnux organization reveals that about 10% of the world's population is affected by communication disorders.

Voice-Based Danger Recognition System uses voice or speech data to detect potential dangers or threats in the context of a voice recording or phone call. This system typically involves the use of Recurrent Neural Network (RNN) deep learning models or Gated Recurrent Neural Networks. The role of the Gated Recurrent Unit (GRU) in a Voice-Based Danger Recognition System is essential, as they enable the system to effectively process, analyze, and interpret voice inputs to detect potential dangers or threats in real-time, ensuring safety and responsiveness in the topic of Speech Emotion Recognition. However, GRU has its incapabilities. According to a recent article on LinkedIn titled "What are the best techniques for using gated recurrent units in machine learning algorithms?" (2024), GRU much like any other neural networks is prone to overfitting where the model becomes overly attuned to the training data, limiting its capacity to generalize new or unseen data. That is why this study is being conducted by the proponents, with the goal to enhance the Gated Recurrent Unit algorithm, to solve the overfitting problem that arises in GRU, and to make its application in Speech Emotion Recognition (SER) more effective in the implementation of Voice-Based Danger Recognition System.

A. Statement of the Problem

Overfitting, which occurs if a model learns to memorize the training data instead of capturing underlying patterns, resulting in poor generalization and high variance to unseen data. This often happens when the model is overly complex or when the training data is limited.

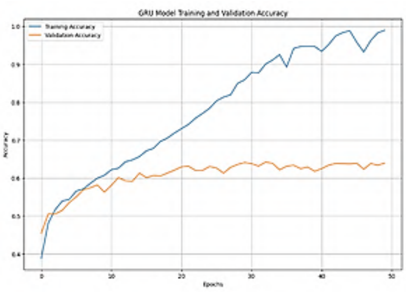

***Figure 1:** shows the Training and Validation Accuracy of the base model trained with an epoch of 50, Batch Size of 64, and 20% of Validation data.*

***Figure 2.** shows the Training and Validation Loss of the base model trained with an epoch of 50, Batch Size of 64, and 20% of Validation data.*

Figures 1 and 2 illustrate the training-validation accuracy and loss of the GRU base model, trained for 50 epochs with a batch size of 64 and 20% validation data to visualize the overfitting problem.

The training accuracy steadily increased from 0.4 to nearly 1.0, while the validation accuracy stayed around 0.6 after the 10th epoch. Similarly, the training loss decreased from 1.5 to nearly 0, whereas the validation loss increased to between 2.5 and 3.0. These results highlight significant overfitting, as the model performs exceptionally well on training data but fails to generalize to validation data

The results below (Figures 3 to 6) are still from the base model trained with an epoch of 50, Batch Size of 64, and 20% of Validation data.

	True Emotion	Predicted Emotion	Predicted Probability
0	disgust	neutral	97.019997
1	neutral	neutral	100.000000
2	disgust	disgust	100.000000
3	sad	sad	100.000000
4	fear	angry	95.639999
...	...	...	...
2950	happy	disgust	51.869999
2951	neutral	fear	98.910004
2952	angry	fear	100.000000
2953	happy	disgust	97.879997
2954	happy	happy	99.610001

Figure 3. Prediction table, which includes 'True Emotion' (actual labels), Predicted Emotion' (model predictions), and 'Predicted Probability' (Confidence of the model on its predictions).

Test Accuracy: 63.21%

***Figure 4.** Test Accuracy*

Volume 06, Issue 03, 2024 | Open Access | ISSN: 2582-6832

Classification Report on test data:

	precision	recall	f1-score	support
fear	0.73	0.67	0.70	527
angry	0.62	0.62	0.62	518
disgust	0.61	0.62	0.62	480
neutral	0.56	0.60	0.58	506
sad	0.63	0.60	0.61	421
happy	0.66	0.68	0.67	503
accuracy			0.63	2955
macro avg	0.63	0.63	0.63	2955
weighted avg	0.63	0.63	0.63	2955

Figure 5. *Classification Report on test data*

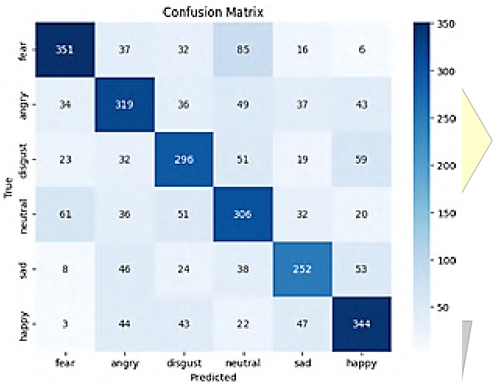

Figure 6. *Confusion Matrix*

Figure 3 presents the Prediction Table, showing True Emotion, Predicted Emotion, and Predicted Probability, used to validate the model's predictions on test data. Figure 4 displays the Test Accuracy, calculated as the percentage of correct predictions where the Predicted Emotion matches the True Emotion. Figure 5 provides the Classification Report, including precision (accuracy of positive predictions), recall (ability to identify positive samples), and the F1 score (harmonic mean of precision and recall). Figure 6 shows the Confusion Matrix, which identifies True Positives (TP), True Negatives (TN), False Positives (FP, Type I error), and False Negatives (FN, Type II error).

B. Objective of the Study

The study primarily aims to create an enhancement of the Gated Recurrent Unit (GRU) algorithm for Speech Emotion Recognition (SER) in the implementation of a Voice-Based Danger Recognition System.

Specifically, it aims to:

1. Accomplish the algorithm to have great generalization and low variance to unseen data. Solving the problem with overfitting makes the models capture and analyze underlying patterns instead of memorizing the training data, this will result in great generalization and low variance to unseen data. With this, overly complex models and even if the training data is limited, overfitting will be resolved for the results.

II. METHODOLOGY

A. Objective

Solve the overfitting problem to make the models capture and analyze underlying patterns instead of memorizing the training data to achieve great generalization and low variance to unseen data.

B. Solution

For the solution, dropout regularization is implemented to help prevent the model from relying too heavily on specific features or patterns in the training data by randomly "dropping out" (i.e., setting to zero) a proportion of the neurons during training. This prevents the algorithm from overfitting and improves its ability to generalize to unseen data. In this model, dropout rate was applied after each GRU layer's computations, including the input gate, update gate, reset gate, candidate memory cell, and final hidden state, ensuring regularization throughout all stages of the algorithm. Batch Normalization was also used to further mitigate the overfitting problem. Batch Normalization is a regularization technique that improves the performance of a deep learning network by first removing the batch mean and then dividing by the batch standard deviation. It was applied after each linear transformation within the GRU computations, including the input gate, update gate, reset gate, and candidate memory cell, before the activation functions (σ and tanh). The goal of batch normalization is to stabilize the training process and improve the generalization ability of the model. By normalizing the activations of a layer, batch normalization reduces overfitting and helps the model generalize better to unseen data.

III. RESULTS AND DISCUSSION

This section will provide the findings of the enhancement of the Gated Recurrent Unit (GRU) algorithm for Speech Emotion Recognition (SER) in the implementation of a Voice-Based Danger Recognition System. In order to mitigate the overfitting problem of the GRU model, Dropout Rate and Batch Normalization were applied. Xavier Glorot Initialization was also used to stabilize the training and convergence. The proponents tried 4 Different variations of 10%, 20%, 30% and 40% Dropout while still incorporating Batch Normalization and Xavier Glorot Initialization. The

United International Journal for Research & Technology
Volume 06, Issue 03, 2024 | Open Access | ISSN: 2582-6832

optimized models were trained the same way as the base model, with an epoch of 50, Batch Size of 64, and 20% of Validation data, in order to keep fairness. A comparative analysis is made to compare the said 4 solutions namely:

1. Dropout Rate of 10%, Batch Normalization and Xavier Glorot
2. Dropout Rate of 20%, Batch Normalization and Xavier Glorot
3. Dropout Rate of 30%, Batch Normalization and Xavier Glorot
4. Dropout Rate of 40%, Batch Normalization and Xavier Glorot

The metrics that were used in order to effectively compare the results of the 4 different solutions are:

1. Test Accuracy
2. Training and Validation Accuracy
3. Training and Validation Loss

Table 1: *Comparative Analysis of the 4 named solutions*

Solutions	**Base Model (No Optimizations)**	**DR10, BN, XG**	**DR20, BN, XG**	**DR30, BN, XG**	**DR40, BN, XG**
Test Accuracy	63.21%	65.65%	67.36%	67.11%	65.28%
Training and Validation Accuracy	Steady Increase of Training Accuracy reaching 1.0 by the last epoch while the Validation Accuracy remains stagnant around 0.6	Close Training and Validation Accuracy with Training Accuracy reaching 0.9 by the end and Validation Data reaching almost 0.7 by the end. but it shows sudden drops on certain epochs on Validation Data.	Closer Training and Validation Accuracy, shows a more stable and consistent trajectory of the Validation Data with Training Accuracy reaching 0.8 and Validation Data reaching almost 0.7 by the last epoch.	Closer Training and Validation Accuracy, with Training Accuracy reaching almost 0.8 and Validation Data reaching almost 0.7 by the last epoch, but shows sudden drops on certain epochs on Validation Data.	Closest Training and Validation Accuracy, shows a more stable and consistent trajectory of the Validation Data with Training Accuracy and Validation Data reaching almost 0.7 by the last epoch.
Training and Validation Loss	Steady decrease of Training Loss reaching almost 0 by the last epoch while the Validation Loss shows steady increase reaching above 2.5 by the last epoch.	Small gap between Training and Validation Loss but shows inconsistency and fluctuations on Validation Data on certain epochs (Noisy Trajectory)	Smaller Gap between Training and Validation Data with more stable and consistent trajectory. (Smooth Trajectory)	Smaller gap between Training and Validation Loss but shows inconsistency and Fluctuations on Validation Data on certain epochs (Noisy Trajectory)	Smallest gap between Training and Validation Data with more stable and consistent trajectory. (Smooth Trajectory)

Legends:
DR*Perc* - Represent Dropout Rate and the rate of the dropout in percentage
BN - Represents Batch Normalization
XG - Represent Xavier Glorot

Shown below are the Test Accuracy, the produced Training-Validation Accuracy and Training-Validation Loss plots of the 4 different solutions in their respective order. For the Test Accuracy, the produced Training-Validation Accuracy and Training-Validation Loss plots of the base model please refer to the figures 1, 2 and 3.

The following Test Accuracy results are from optimized models with Batch Normalization and Xavier Glorot Initialization, they only differ in Dropout rates:

```
Test Accuracy: 65.65%
```

Figure 7. *with Dropout Rate of 10%*

```
Test Accuracy: 67.34%
```

Figure 8. *with Dropout Rate of 20%*

```
Test Accuracy: 67.11%
```

Figure 9. *with Dropout Rate of 30%*

Volume 06, Issue 03, 2024 | Open Access | ISSN: 2582-6832

```
Test Accuracy: 65.28%
```

Figure 10. *with Dropout Rate of 40%*

As shown in the figures above, the optimized model with Dropout Rate of 20%, Batch Normalization and Xavier Glorot was able to produce the highest improvement in terms of Test Accuracy with 67.34%. Compared to the base model with no optimizations that was able to produce 63.21% of Test Accuracy, the optimized model with Dropout Rate of 20%, Batch Normalization and Xavier Glorot was able to improve it by 4.13%. This proves that the solution was able to help improve the metric of Test Accuracy.

The following figures will show produced Training-Validation Accuracy of the optimized model with Batch Normalization and Xavier Glorot Initialization, they only differ in Dropout rates:

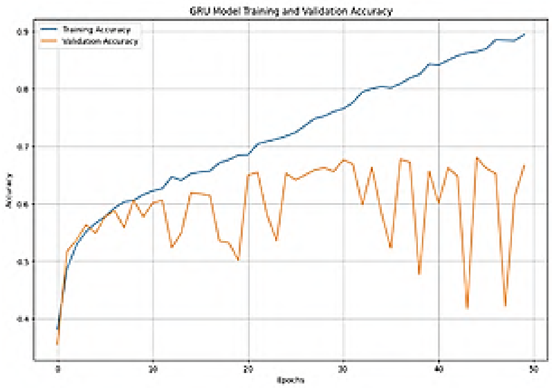

Figure 11. *shows the Training and Validation Accuracy of the Optimized model with a Dropout Rate of 10%*

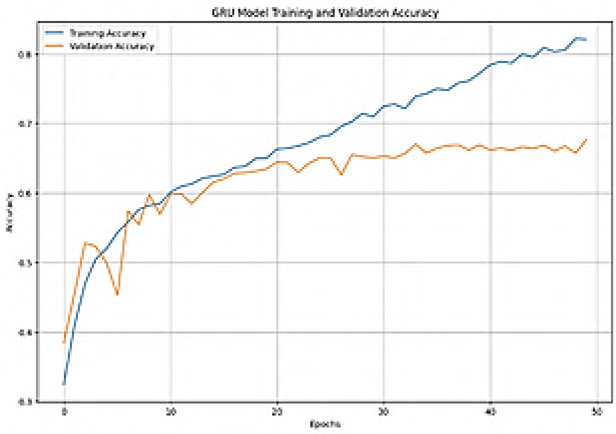

Figure 12. *shows the Training and Validation Accuracy of the Optimized model with a Dropout Rate of 20%*

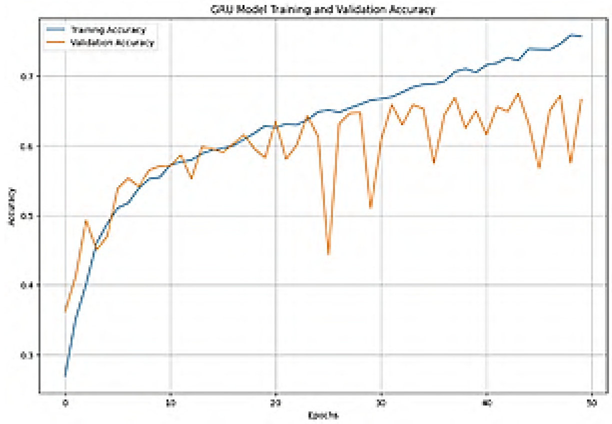

Figure 13. *shows the Training and Validation Accuracy of the Optimized model with a Dropout Rate of 30%*

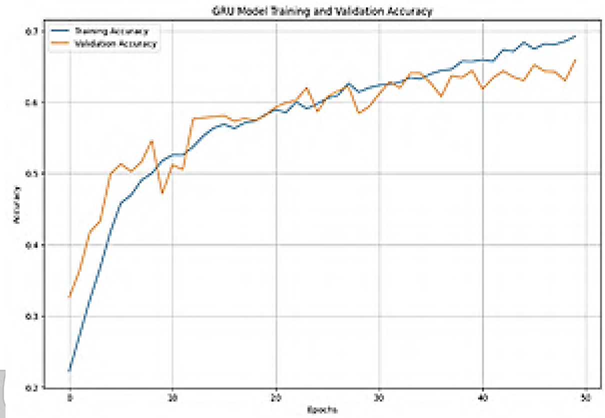

Figure 14. *shows the Training and Validation Accuracy of the Optimized model with a Dropout Rate of 40%*

Figures 11-14 show the differences between the 4 optimized models with Batch Normalization and Xavier Glorot Initialization in terms of Training-Validation Accuracy. The optimized model with a Dropout Rate of 10% was able to improve the gap between the Training and Validation Accuracy compared to the base model but the Validation Accuracy is inconsistent and suddenly drops on certain epochs while the Training Accuracy steadily increases. The optimized model with a Dropout Rate of 20, on the other hand, shows a much greater improvement. It shows a steady increase on both Training and Validation Accuracy with the Training Accuracy reaching 0.8 and Validation Accuracy reaching almost 0.7 by the last epoch. The optimized model with a Dropout Rate of 30% showed a similar but still better results to Figure 11 where the gap between the Training and Validation Accuracy compared to the base model is smaller but the Validation Accuracy is inconsistent and suddenly drops on certain epochs.

UIJRT ISSN: 2582-6832
United International Journal for Research & Technology
Volume 06, Issue 03, 2024 | Open Access | ISSN: 2582-6832

Lastly, the optimized model with a Dropout Rate of 40% showed the best improvement among all the solutions in terms of Training and Validation Accuracy with both steadily increasing while being the closest to each other.

The following figures will show produced Training-Validation Loss of the optimized model with Batch Normalization and Xavier Glorot Initialization, they only differ in Dropout rates:

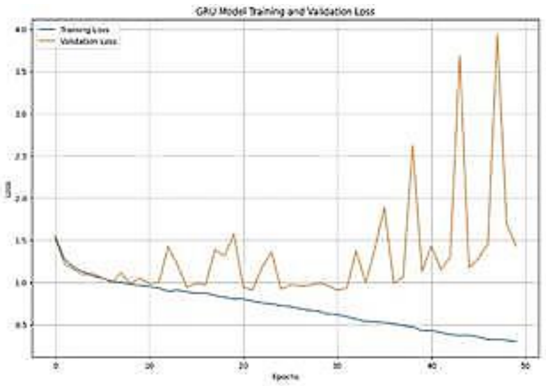

Figure 15. *shows the Training and Validation Loss of the Optimized model with a Dropout Rate of 10%*

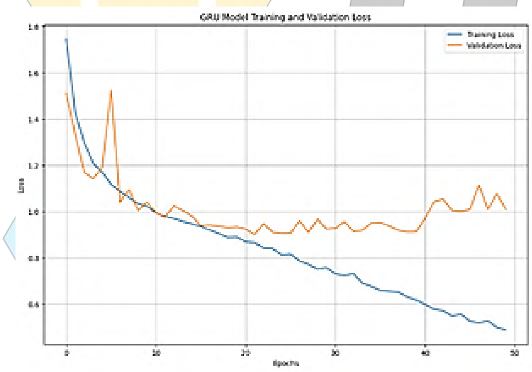

Figure 16. *shows the Training and Validation Loss of the Optimized model with a Dropout Rate of 20%*

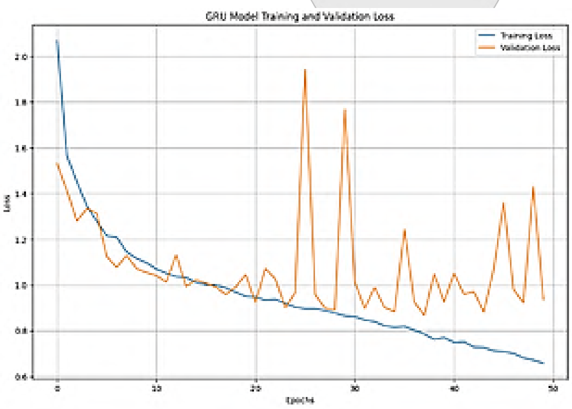

Figure 17. *shows the Training and Validation Loss of the Optimized model with a Dropout Rate of 30%*

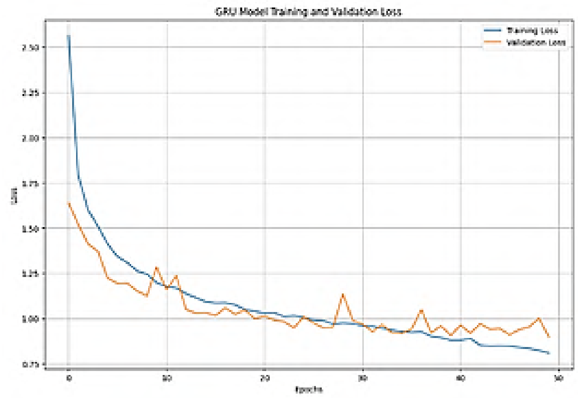

Figure 18. *shows the Training and Validation Loss of the Optimized model with a Dropout Rate of 40%*

Figures 15 - 18 shows the Training and Validation Loss of the 4 optimized models. Figure 15 shows improvement compared to the base model in terms of Training and Validation Loss. It shows a small gap between Training and Validation Loss but shows inconsistency and fluctuations on Validation Data on certain epochs.

Figure 16 shows an even greater improvement in terms of Training and Validation Loss. It was able to make a smaller gap between Training and Validation Data with a more stable and consistent trajectory.

Figure 17, on the other hand, although was able to show improvement from the base model suffers from fluctuating and inconsistent Validation Loss much like on Figure 15. Lastly, figure 18 shows the best improvement in terms of Training and Validation Loss. It was able to closely tighten the gap between the Training and Validation Loss with a more stable and consistent trajectory.

Considering the results of the 4 optimized models and after careful deliberation, the optimized model with a Dropout rate of 20%, Batch Normalization and Xavier Glorot was deemed the best solution out of all the 4 solutions.

Although the model with Dropout rate of 40%, Batch Normalization and Xavier Glorot showed a better Training-Validation Accuracy and Loss, its performance on Test Data is relatively lower compared to the optimized model with a Dropout rate of 20%, Batch Normalization and Xavier Glorot, and actually has the lowest Test Accuracy among all the 4 solutions.

The optimized model with a Dropout rate of 20%, Batch Normalization and Xavier Glorot was able to mitigate the overfitting problem shown through the improvement in terms of Training-Validation Accuracy and Loss compared to the base model, without compromising the Test Accuracy and improving it effectively by increasing it by 4.13%, thus making it the best solution out of all the 4 solutions.

To further solidify that the optimized model with a Dropout rate of 20%, Batch Normalization and Xavier Glorot was able to improve the base model, the proponents also compared their Classification Reports and Confusion Matrices.

For the Classification Report of the base model refer to Figure 5 and for the Confusion Matrix refer to Figure 6.

Classification Report on test data:

	precision	recall	f1-score	support
fear	0.75	0.76	0.75	527
angry	0.66	0.62	0.64	518
disgust	0.73	0.64	0.68	480
neutral	0.58	0.70	0.63	506
sad	0.62	0.65	0.64	421
happy	0.72	0.68	0.70	503
accuracy			0.67	2955
macro avg	0.68	0.67	0.67	2955
weighted avg	0.68	0.67	0.67	2955

***Figure 19.** shows the Classification Report on test data of the Optimized model with a Dropout Rate of 20%, Batch Normalization and Xavier Glorot Initialization.*

The figure above shows that the optimized model with a Dropout rate of 20%, Batch Normalization and Xavier Glorot, outperformed the base model across all metrics.

The optimized model achieved an accuracy of 67%, compared to 63% for the base model.

Additionally, the macro and weighted averages for precision, recall, and F1-score improved from 0.63 in the base model to 0.68, 0.67, and 0.67 respectively in the optimized model.

Class-wise performance also saw notable improvements, with all F1-scores increasing.

Specifically, the fear class improved from 0.70 to 0.75, and disgust rose from 0.62 to 0.68. Other classes, such as neutral and happy, also showed consistent gains.

These results highlight the effectiveness of the applied optimizations in enhancing the model's performance.

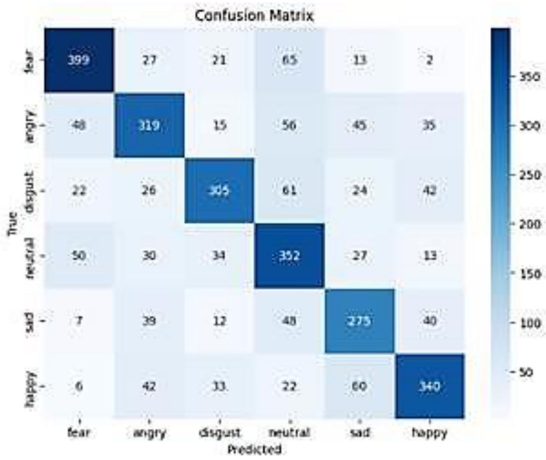

***Figure 20.** shows the Confusion Matrix of the Optimized model with a Dropout Rate of 20%, Batch Normalization and Xavier Glorot Initialization.*

The figure above shows that the optimized model with a Dropout rate of 20%, Batch Normalization and Xavier Glorot performs better overall compared to the base model, as it improves accuracy for key classes such as fear, neutral, and sad. Specifically, the optimized model correctly predicts 399 instances of fear (up from 351), 352 instances of neutral (up from 306), and 275 instances of sad (up from 252). While the angry class remains consistent across both models, the happy class sees a slight drop in correct predictions from 344 to 340, which is negligible compared to the significant gains in other categories. Additionally, the optimized model reduces misclassifications for critical classes like fear and neutral, leading to a more balanced performance overall. Given these improvements, the optimized GRU model is the better-performing model.

IV. CONCLUSION

With the optimization and enhancements applied, the problem with overfitting, which makes the models memorize the training data instead of capturing patterns is minimized. The findings demonstrated that applying Dropout Rate of 20%, Batch Normalization, and Xavier/Glorot Initialization in the Gated Recurrent Unit model notably influenced the test accuracy, training and validation accuracy, training and validation loss. The results showed a smaller gap between training and validation accuracy, which indicates a more stable and consistent performance. By the final epoch, the training accuracy reaches 0.8, while the validation accuracy approaches 0.7, reflecting a stable trajectory for both datasets. The model with the applied optimizations showed a depreciation of the overfitting problem compared to the base model and with other optimization techniques, which was exhibited and shown in the increase of test accuracy by 4.13%. These results provide a deeper understanding of Gated Recurrent Unit and offer an enhanced version for further implementations aimed at the application in the field of speech emotion recognition. This study contributes to the field of deep learning applied in real-world systems, which benefit a large portion of the population of the globe.

V. RECOMMENDATIONS

Future research could explore the enhanced version of gated recurrent unit in greater depth with the suggestions of the researchers. The first recommendation is to apply the optimizations one at a time, which means using a single technique among Dropout Rate of 20%, Batch Normalization, and Xavier Glorot Initialization to see the variations of the results. The second recommendation is to apply data augmentation into the dataset to tweak and modify the diversity of the data, which can impact the training and validation accuracy. The last recommendation, which is the most crucial among the suggestions, is testing the enhanced model in other types of data, other than speech and audio data. With that recommendation also comes an implementation of the future models to various types of systems that encompass fields near and far the field of speech emotion recognition.

REFERENCES

[1] Al-Selwi, A., et al. (n.d.). RNN-LSTM: From applications to modeling techniques and beyond—Systematic review. Journal of King Saud University - Computer and Information Sciences. https://www.sciencedirect.com/science/article/pii/S1319157824001575

[2] Analytics Vidhya. (2021). Introduction to gated recurrent unit (GRU). Retrieved from https://www.analyticsvidhya.com/blog/2021/03/introduction-to-gated-recurrent-unit-gru/

[3] Diao, E., Ding, J., & Tarokh, V. (2019). Restricted recurrent neural networks. IEEE. arxiv.org/pdf/1908.07724

[4] Desai, C. (n.d.). Impact of weight initialization techniques on neural network efficiency and performance: A case study with MNIST dataset. International Journal of Engineering and Computer Science. https://www.researchgate.net/publication/379696245_Impact_of_Weight_Initialization_Techniques_on_Neural_Network_Efficiency_and_Performance_A_Case_Study_with_MNIST_Dataset

[5] Educative. (2023). What is a gated recurrent unit (GRU)? Retrieved from https://www.educative.io/answers/what-is-a-gated-recurrent-unit-gru

[6] GeeksforGeeks. (2023). Gated recurrent unit networks. Retrieved from https://www.geeksforgeeks.org/gated-recurrent-unit-networks/

[7] Kelleher, J. D. (2019). Deep learning. MIT Press. https://direct.mit.edu/books/book/4556/Deep-Learning

[8] Liu, G., Cai, S., & Wang, C. (2023). Speech emotion recognition based on emotion perception. EURASIP Journal on Audio, Speech, and Music Processing. https://asmp-eurasipjournals.springeropen.com/articles/10.1186/s13636-023-00289-4

[9] Medium (Anishnama). (2023). Understanding gated recurrent unit (GRU) in deep learning. Retrieved from https://medium.com/@anishnama20/understanding-gated-recurrent-unit-gru-in-deep-learning-2e54923f3e2

[10] Medium (Shevchenko, E.). (2024). Regularization methods for recurrent networks. Retrieved from https://medium.com/@eugenesh4work/regularization-methods-for-recurrent-networks-215e0147d922#:~:text=Dropout%20in%20Recurrent%20Neural%20Networks&text=Dropout%20involves%20randomly%20setting%20a,heavily%20on%20any%20one%20feature

[11] Medium (Sundararajan, B.). (2024). Dropout: The secret weapon for regularizing neural networks. Retrieved https://medium.com/@bragadeeshs/dropout-the-secret-weapon-for-regularizing-neural-networks-338f079ebc29

[12] Medium (Olamendy, J. C.). (2024). Weight initialization for deep learning neural networks. Retrieved from https://medium.com/@juanc.olamendy/weight-initialization-for-deep-learning-neural-networks-6047cbe27297

[13] Onwujekwe, G., Haber, B., & Bajaj, V. (2024). Using gated recurrent unit (GRU) for speech emotion recognition in children. International Journal of Science, Engineering and Management (IJSEM). https://ijsem.org/article/6.pdf

[14] Prakash, P. R., et al. (2022). A novel convolutional neural network with gated recurrent unit for automated speech emotion recognition and classification. Journal of Control and Decision. https://www.tandfonline.com/doi/abs/10.1080/23307706.2022.2085198

[15] Salehin, I., & Kang, D. K. (2023). A review on dropout regularization approaches for deep neural networks within the scholarly domain. Electronics Journal. https://www.researchgate.net/publication/372410036_A_Review_on_Dropout_Regularization_Approaches_for_Deep_Neural_Networks_within_the_Scholarly_Domain

[16] Shanthamallu, U. S., & Spanias, A. (2022). Machine and deep learning algorithms and applications. Springer Nature. https://link.springer.com/book/10.1007/978-3-031-03758-0

[17] Stanford, A. W., & Matsukawa, A. (n.d.). Exploring batch normalization in recurrent neural networks. Stanford. jaywhang.com/assets/batchnorm_rnn.pdf

[18] Yan, Y., & Shen, X. (2022). Research on speech emotion recognition based on AA-CBGRU network. Shanghai Institute of Technology. https://www.mdpi.com/2079-9292/11/9/1409

[19] Yuan, J., & Tian, Y. (2019). An intelligent fault diagnosis method using GRU neural network towards sequential data in dynamic processes. MDPI. https://www.mdpi.com/2227-9717/7/3/152

[20] de Lope, J., & Graña, M. (2023). An ongoing review of speech emotion recognition. Elsevier B.V.. https://www.sciencedirect.com/science/article/pii/S0925231223000103

[21] LinkedIn. (2024). What are the best techniques for using gated recurrent units in machine learning algorithms? Retrieved from https://www.linkedin.com/advice/1/what-best-techniques-using-gated-recurrent-units-dslce