Model Card: ResNet18

1 Model Overview

Architecture: ResNet18 employs residual (skip) connections, which help mitigate the vanishing gradient problem during training of deep networks.

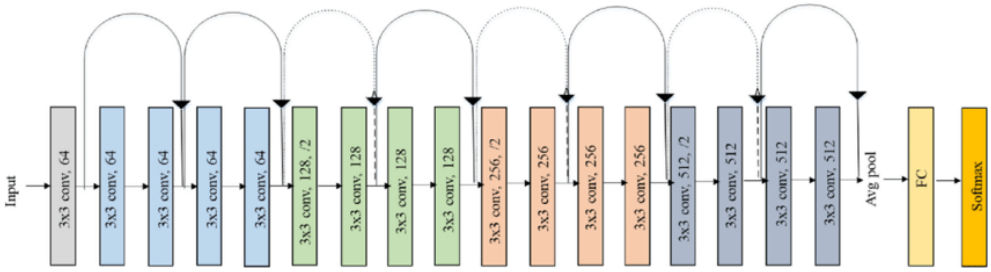

Figure 1: ResNet18 Architecture [1]

Training Data: The model is trained on the ASL Alphabet dataset containing 87,000 images resized to 200x200 pixels, spanning 29 classes (26 letters plus 3 additional classes to aid live classification). 10 percent is held out as a test set, so 78,300 images are used in training.

Table 1: Overview of Data Split and Image Specifications

Category	Number of Images	Percentage
Overall Dataset	87,000	100%
Training Set	78,300	90%
Training (90%)	62,640	72%
Validation (10%)	15,660	18%
Test Set	8,700	10%

Hyperparameters:

- Optimizer: Adam
- Loss Function: Cross-Entropy
- Learning Rate: 0.01
- Batch Size: 64
- Epochs: 20 (best validation accuracy selected)

2 Intended Use

ResNet18 is an instance of the ResNet architecture designed for image classification. Our ResNet18 implementation uses a pre-trained ResNet18 model with transfer learning to classify images of American Sign Language (ASL) letters. The model is intended for:

- Classifying American Sign Language alphabet letters from images
- Real-time sign language interpretation using live video inputs

3 Performance

On a separate test dataset, the ResNet18 model achieved:

- **Test Accuracy:** 100%
- **Precision:** 1.0
- **Recall:** 1.0

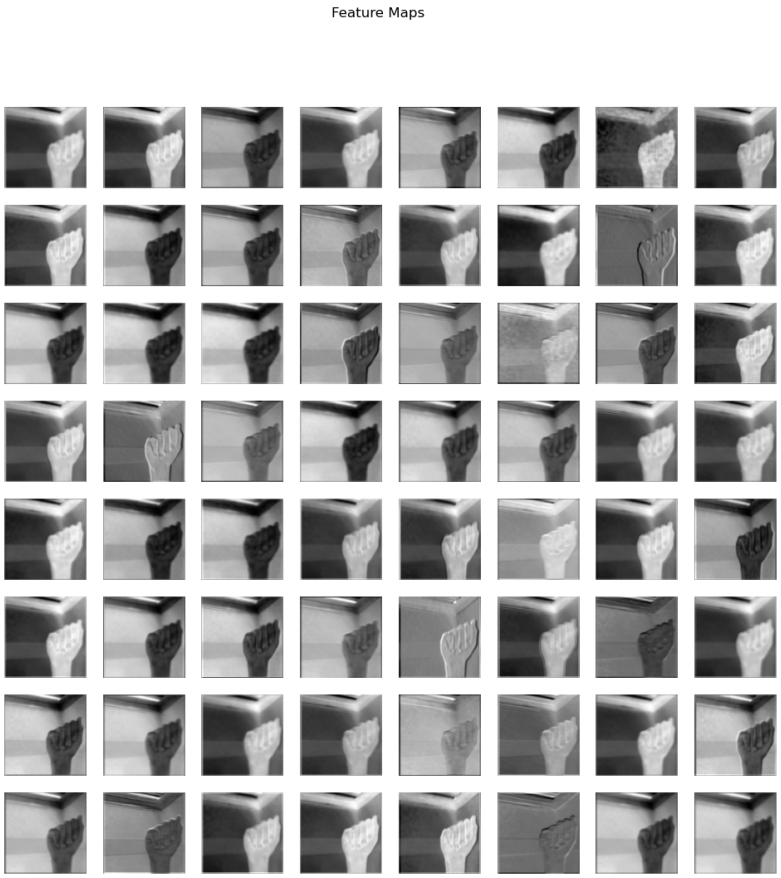

Figure 2: Feature Maps from the ResNet18 model

Figure 2 shows sample feature maps extracted from intermediate layers of the model.

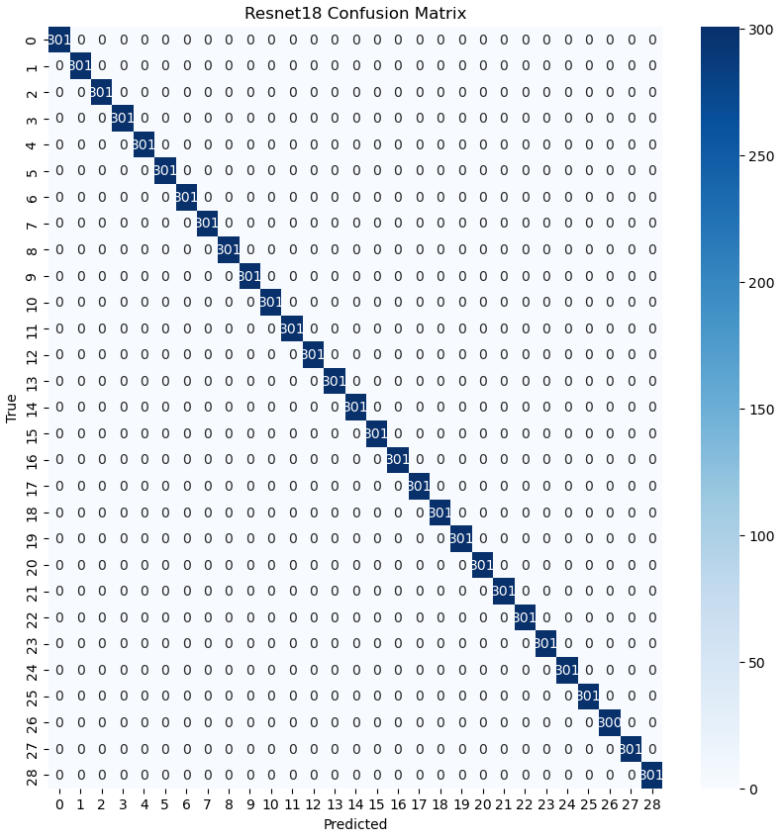

Figure 3: Confusion Matrix for ResNet18

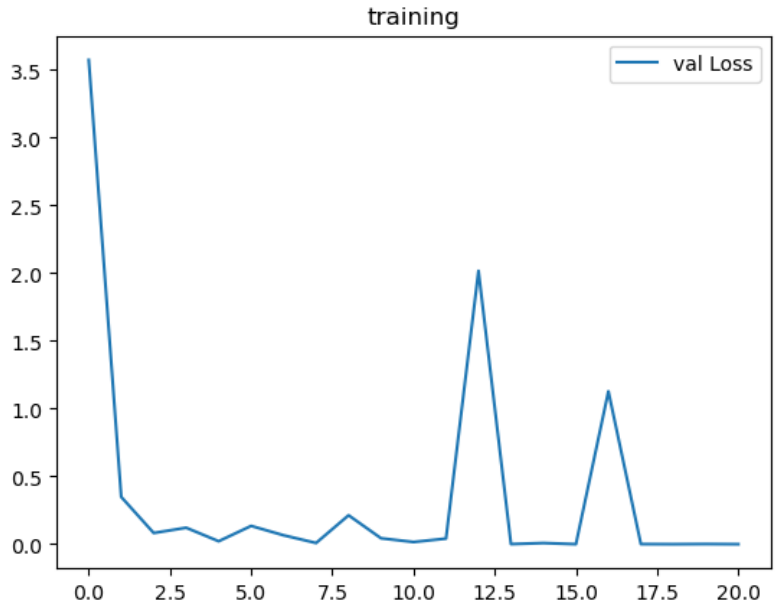

Figure 4: Training Loss for ResNet18

4 Limitations

Despite high accuracy, some limitations are noted:

- Sensitivity to variations in image quality and lighting conditions during live classification
- Dependence on precise hand positioning for real-time classification
- Potential challenges in deployment on resource-constrained devices

5 Ethical Considerations

- Ensure the ASL dataset is diverse and representative to avoid biased outcomes
- Implement privacy protections when deploying real-time video classification

References

[1] "A deep learning approach for automated diagnosis and multi-class classification of alzheimer's disease stages using resting-state fmri

and residual neural networks," https://www.researchgate.net/publication/336642248_A_Deep_Learning_Approach_for_Automated_Diagnosis_and_Multi-Class_Classification_of_Alzheimer's_Disease_Stages_Using_Resting-State_fMRI_and_Residual_Neural_Networks, accessed: 2025-05-09.