**Report: RSNA Pneumonia Detection Model (V30)**

**Objective**: Develop a robust deep learning model to classify chest X-rays as **Normal** or **Pneumonia** using the RSNA dataset, addressing **overfitting**, **class imbalance**, and **data pipeline instability**.

**Dataset Overview**

* **Source**: RSNA Pneumonia Detection Challenge.
* **Class Distribution**:
  + **Normal**: ~75% (4135 samples).
  + **Pneumonia**: ~25% (1282 samples). <!-- Updated support count -->
* **Image Resolution**: Resized to **224x224** (DenseNet compatibility).

**Model Architecture (V30)**

**Base Model: DenseNet201 (pretrained on ImageNet).**

* **Input Shape**: (224, 224, 3).
* **Fine-Tuning**:
  + **Frozen Layers**: First 200 layers.
  + **Trainable Layers**: Last 53 layers.

**Custom Head:**

python

Copy

Sequential([

SpatialDropout2D(0.7), # Increased from 0.5

GlobalAveragePooling2D(),

Dense(64, activation='relu', kernel\_regularizer=tf.keras.regularizers.l2(0.05)), # L2 increased from 0.01

Dropout(0.6), # Increased from 0.5

Dense(1, activation='sigmoid')

])

**Loss & Optimizer:**

* **Loss**: BinaryFocalCrossentropy(gamma=3.0) → Adjusted from gamma=2.0.
* **Optimizer**: AdamW(learning\_rate=1e-5, clipnorm=1.0) → Reduced LR + gradient clipping.
* **Metrics**: Accuracy, AUC.

**Training Configuration**

* **Batch Size**: 32.
* **Epochs**: 30 (early stopping with patience=3).
* **Class Weights**: {0: 1, 1: 8} → Increased penalty for Pneumonia.
* **Augmentation**:
  + Random flips, brightness, rotation (±20°), and crops (180x180 → 224x224).
* **Data Pipeline Fixes**:
  + Added .repeat() to training data for infinite batches.
  + Recalculated steps\_per\_epoch = len(train\_df) // BATCH\_SIZE.

**Results (Post-Fixes)**

| **Metric** | **Train** | **Validation** |
| --- | --- | --- |
| **AUC** | 0.88 | 0.87 |
| **Loss** | 0.22 | 0.17 |

**Classification Report:**

| **Class** | **Precision** | **Recall** | **F1-Score** | **Support** |
| --- | --- | --- | --- | --- |
| **Normal** | 0.94 | 0.75 | 0.83 | 4135 |
| **Pneumonia** | 0.49 | 0.84 | 0.62 | 1282 |
| **Accuracy** | - | 0.77 | - | 5337 |

**Key Observations:**

1. **Improved Stability**: Reduced erratic accuracy jumps after pipeline fixes.
2. **Validation AUC**: Reached **0.87** (up from 0.60 pre-fix).
3. **Class Imbalance**: High Pneumonia recall (84%) but low precision (49%) persists.

**Key Fixes in V30**

1. **Overfitting Mitigation**:
   * Increased **SpatialDropout2D(0.7)** and **Dropout(0.6)**.
   * Strengthened **L2 regularization (0.05)**.
2. **Class Imbalance**:
   * Adjusted **class weights to {0:1, 1:8}** and **focal loss (gamma=3.0)**.
3. **Training Stability**:
   * Added gradient clipping (clipnorm=1.0).
   * Reduced learning rate to **1e-5**.
4. **Data Pipeline**:
   * Fixed infinite batching with .repeat().
   * Corrected steps\_per\_epoch.

**Remaining Challenges**

1. **Low Pneumonia Precision**: Too many false positives.
2. **Validation AUC Plateau**: Stagnant at 0.87 despite training improvements.

**Next Steps**

1. **Hyperparameter Tuning**: Optimize dropout, L2, and focal loss gamma with Keras Tuner.
2. **Test-Time Augmentation (TTA)**: Improve precision via multi-crop inference.
3. **Error Analysis**: Visualize false positives using Grad-CAM.

**Conclusion**

**V30** significantly improved stability and validation performance (AUC **0.87** vs. initial 0.60) but requires further tuning to address precision and AUC plateauing. Future work will focus on **TTA** and **advanced regularization** to reduce false positives.