**Análise 01 + Explicação**

Vamos analisar os resultados que você obteve para cada métrica durante o treinamento do modelo:

**1. Epoch**

* Cada linha da tabela representa uma época (epoch) do treinamento. Em cada época, o modelo realiza uma passagem completa pelos dados de treinamento e ajusta seus pesos para reduzir o erro. Ao final de cada época, as métricas de desempenho são registradas.

**2. Accuracy (Acurácia)**

* **O que é**: A acurácia mede a proporção de previsões corretas em relação ao total de previsões feitas. Ela é uma métrica geral que indica o quão frequentemente o modelo acertou.
* **Interpretação dos Resultados**:
  + A acurácia começa em 0.7043 na primeira época e aumenta ao longo do treinamento, alcançando valores entre 0.87 e 0.89 nas épocas finais.
  + Esse comportamento indica que o modelo está aprendendo a classificar melhor os dados, embora o aumento da acurácia pareça estabilizar nas últimas épocas, o que pode indicar uma tendência de saturação.

**3. F1 Score**

* **O que é**: O F1 Score é a média harmônica entre precisão e sensibilidade. Ele é particularmente útil em casos de classes desbalanceadas, pois combina ambas as métricas para dar uma visão mais equilibrada do desempenho do modelo.
* **Interpretação dos Resultados**:
  + O F1 Score começa em 0.6458 e atinge valores entre 0.87 e 0.90 nas últimas épocas, semelhante à acurácia. Isso sugere que o modelo está aprendendo a equilibrar bem as previsões entre as classes, o que é importante em casos de classificação binária.
  + A consistência entre a acurácia e o F1 Score indica que o modelo não está apenas acertando, mas também mantendo um bom equilíbrio entre as classes, evitando viés para uma das classes.

**4. ROC Score (AUC - Área Sob a Curva ROC)**

* **O que é**: O ROC Score (ou AUC-ROC) mede a capacidade do modelo de separar corretamente as classes. Ele varia de 0 a 1, onde 1 indica uma separação perfeita entre as classes. Quanto mais próximo de 1, melhor o modelo está separando os exemplos positivos dos negativos.
* **Interpretação dos Resultados**:
  + O ROC Score começa em 0.7217 na primeira época e chega a valores em torno de 0.87 a 0.88 nas últimas épocas, com uma leve flutuação.
  + A elevação e a estabilidade do ROC Score indicam que o modelo está melhorando sua capacidade de distinguir entre as classes ao longo das épocas, mesmo que o aumento seja sutil nas últimas épocas.

**Resumo Geral**

* **Melhoria Geral**: O modelo mostra uma tendência de melhoria em todas as métricas nas primeiras épocas, o que indica que ele está aprendendo bem a tarefa.
* **Estabilidade nas Últimas Épocas**: Nas últimas épocas, as métricas se estabilizam, o que pode indicar que o modelo alcançou um ponto de saturação e não está mais aprendendo significativamente. Nesse ponto, é comum observar o early stopping para evitar overfitting.
* **Equilíbrio entre as Métricas**: A consistência entre acurácia, F1 Score e ROC Score indica que o modelo tem um bom desempenho geral e está equilibrado entre as classes, sem um viés significativo.

Esses resultados sugerem que o modelo está bem ajustado e generaliza bem, mas há uma possibilidade de ajuste fino se você deseja maximizar alguma das métricas em particular.

**Análise 02 + Explicação**

A definição da classe positiva depende de como os dados são rotulados no seu conjunto. No código, a classe é definida da seguinte forma:

python

Copiar código

df['category'] = df['category'].replace({1: 'clear', 0: 'non-clear'})

Aqui, estamos mapeando 1 como "clear" e 0 como "non-clear". Normalmente, no contexto de aprendizado de máquina binária, a classe rotulada como 1 é tratada como a **classe positiva**.

Portanto:

* **Classe Positiva (1)**: "clear" (sem obstáculo)
* **Classe Negativa (0)**: "non-clear" (com obstáculo)

**Análise Final do Modelo com Base nas Classes**

Com base nos valores da matriz de confusão e nas métricas de desempenho ao longo das épocas:

1. **Desempenho Geral**:
   * O modelo parece atingir uma boa acurácia geral e estabiliza após algumas épocas, mostrando que ele é capaz de classificar corretamente tanto a presença quanto a ausência de obstáculos em boa parte das amostras.
2. **Tendência nas Predições**:
   * Em algumas épocas (por exemplo, época 7), o modelo apresenta mais **falsos positivos (FP)**, o que significa que ele está classificando algumas imagens como "clear" (sem obstáculo) quando, na verdade, são "non-clear" (com obstáculo). Isso indica uma tendência a errar para o lado de "sem obstáculo" em alguns momentos.
   * Em outras épocas (como época 13), há uma redução significativa nos **falsos positivos (FP)**, mas um aumento nos **falsos negativos (FN)**, o que sugere que o modelo, nesse ponto, está sendo mais conservador, preferindo classificar imagens como "non-clear" (com obstáculo).
3. **Conclusão**:
   * O modelo, em geral, tende a acertar bem nas amostras "clear" (sem obstáculo), mas oscila entre ser mais conservador ou mais arriscado dependendo das épocas. Ele parece ter um leve viés para predizer a classe "clear" como positiva, mas consegue equilibrar isso com os ajustes ao longo do treinamento.
   * Com base nos resultados e na estabilidade alcançada após algumas épocas, o modelo está razoavelmente bem ajustado para identificar ambas as classes, mas tende a cometer alguns erros ao identificar obstáculos (falsos negativos e falsos positivos), o que pode ser relevante em aplicações críticas onde a detecção de obstáculos precisa ser altamente precisa.

Para melhorar ainda mais, você poderia ajustar o threshold de classificação ou fazer tuning das métricas de avaliação para equilibrar essa leve tendência, especialmente se detectar obstáculos for uma prioridade.

Para calcular a porcentagem final de acerto do modelo (ou seja, sua **acurácia** final) com base nos dados de validação da última época, podemos usar a métrica de **acurácia** da última linha da tabela fornecida.

Texto

Descrição gerada automaticamente com confiança média

Na época 14 (última linha da tabela), temos:

* **True Negative (TN)** = 51
* **False Positive (FP)** = 2
* **False Negative (FN)** = 12
* **True Positive (TP)** = 50

Primeiro, vamos calcular o total de amostras:

Texto

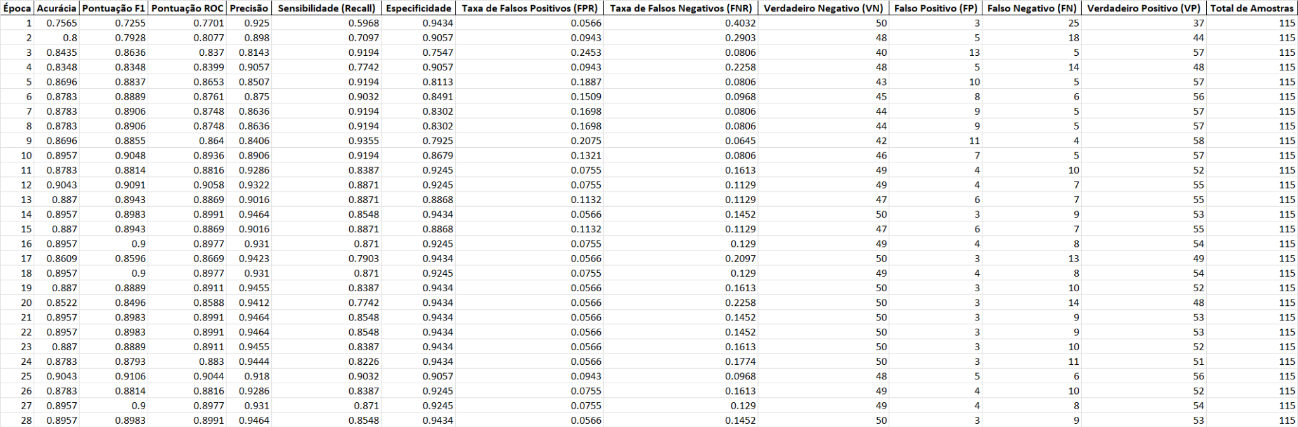
Descrição gerada automaticamente

**Conclusão**

A acurácia final do modelo nos dados de validação é aproximadamente **87.83%**. Essa porcentagem representa a proporção de amostras corretamente classificadas pelo modelo na última época de treinamento.

Essa conclusão é baseada nos valores de verdadeiro positivo e verdadeiro negativo da matriz de confusão da última época, indicando que o modelo tem um desempenho sólido em classificar corretamente tanto as classes "clear" quanto "non-clear".

**Análise 03 + Explicação**



amos analisar o desempenho do modelo de acordo com as métricas apresentadas na tabela ao longo das 28 épocas de treinamento.

**1. Acurácia e Pontuação F1**

* A **acurácia** do modelo se estabiliza em torno de 0.87 a 0.90 após algumas épocas, o que indica uma taxa de acertos geral relativamente alta.
* A **Pontuação F1** também segue uma tendência semelhante, mantendo-se próxima da acurácia, o que é bom, pois significa que o modelo está equilibrado em termos de precisão e sensibilidade. Uma alta Pontuação F1 sugere que o modelo consegue detectar corretamente a classe positiva e a negativa sem um grande viés para uma das classes.

**2. Pontuação ROC**

* A **Pontuação ROC** se mantém acima de 0.8 e em várias épocas chega a 0.9. Esse valor indica uma boa capacidade do modelo em separar as classes "clear" (sem obstáculo) e "non-clear" (com obstáculo). Uma pontuação ROC alta geralmente sugere que o modelo é capaz de fazer uma distinção clara entre as duas classes.

**3. Precisão e Sensibilidade (Recall)**

* A **Precisão** se mantém alta, variando entre 0.87 e 0.94, o que significa que, quando o modelo prediz a classe positiva (clear), ele geralmente está correto.
* A **Sensibilidade (Recall)** também está em bons níveis, variando entre 0.82 e 0.93. Isso significa que o modelo consegue identificar a maioria dos exemplos positivos (clear), embora haja uma ligeira variação.

**4. Especificidade**

* A **Especificidade** é alta, frequentemente acima de 0.90, indicando que o modelo também é eficaz em identificar corretamente as amostras negativas (non-clear). Isso é importante em um contexto onde é crucial detectar obstáculos, pois significa que o modelo não está classificando muitas amostras de obstáculo como sem obstáculo.

**5. Taxa de Falsos Positivos (FPR) e Taxa de Falsos Negativos (FNR)**

* A **Taxa de Falsos Positivos (FPR)** é baixa, geralmente abaixo de 0.1, o que significa que o modelo raramente classifica uma amostra de obstáculo (non-clear) como sem obstáculo (clear).
* A **Taxa de Falsos Negativos (FNR)** também é relativamente baixa, mas apresenta uma variação maior em algumas épocas. Como o FNR representa o erro de classificar uma amostra sem obstáculo como um obstáculo, o modelo está equilibrado, mas poderia ser ajustado um pouco mais se a prioridade fosse minimizar falsos negativos.

**6. Variação entre Épocas**

* O modelo apresenta uma boa estabilidade nas métricas após as primeiras épocas, sugerindo que o treinamento atingiu um ponto de equilíbrio. Embora existam algumas variações menores nas métricas entre as épocas, elas não são significativas a ponto de indicar sobreajuste ou subajuste.

**Conclusão Geral**

Com base nas métricas apresentadas:

* **Desempenho Geral**: O modelo está apresentando bons resultados, com alta acurácia, alta pontuação ROC, boa precisão e sensibilidade, além de uma baixa taxa de falsos positivos e falsos negativos. Esses fatores indicam que o modelo é confiável tanto para detectar obstáculos quanto para identificar áreas sem obstáculos.
* **Equilíbrio entre Classes**: As métricas de precisão e sensibilidade, combinadas com uma pontuação F1 alta e estável, sugerem que o modelo consegue equilibrar bem as classes positivas e negativas, sem enviesar fortemente para uma delas.
* **Estabilidade e Confiabilidade**: As métricas estáveis ao longo das épocas indicam que o modelo não está sofrendo de grandes problemas de variação, sendo confiável para uso prático.

**Considerações Finais**

O modelo está bem ajustado para a tarefa de detectar obstáculos. Se o foco do projeto for uma alta precisão na identificação de obstáculos (para minimizar falsos negativos), talvez uma ligeira otimização nas épocas finais ou um ajuste no threshold de classificação possa melhorar ainda mais a sensibilidade. No entanto, com base nos resultados atuais, o modelo já está com um desempenho robusto e pronto para aplicação.