The background features a dark blue field with intricate white and light blue circuit traces, nodes, and gear-like shapes. A prominent cyan square is located in the bottom-left corner, and a series of vertical bars, some solid cyan and some outlined, are positioned in the bottom-right.

UM ESTUDO SOBRE SEGMENTAÇÃO DE IMAGENS COM REDES CONVOLUCIONAIS

O problema da segmentação de texto em mangá

Pedro H. B. de Almeida

Estudante

Nina S. T. Hirata

Orientadora

The collage consists of several manga panels:

- Top Left Panel:** A character with long dark hair looks distressed. Text bubbles say: "今部訓装置で起るんだって... 同じようなの!!" (It's starting with the current training device... it's the same kind!!).
- Top Middle Panel:** A close-up of a control panel with many buttons and dials. Text says: "おやつとこれどうなってるの? 計器が.....!" (What about snacks and this? The gauges are.....!).
- Top Right Panel:** A character with glasses and a headband looks frustrated while holding a controller. Large bold text reads: "!! ダメですツツ 推力姿勢制御とも不能!!" (!! It's no good Tsutsu, neither thrust attitude control nor anything else is working!!). Below him, another character says: "反応しませんよッ" (It doesn't respond).
- Middle Left Panel:** A night view of a city with a bridge. A speech bubble says: "N向かうほうに杉田ワッってことだ" (That means going towards the direction where Sugiwa is, right?).
- Middle Right Panel:** A view of a large building at night.
- Bottom Left Panel:** A dramatic scene of a ship flying through a dark, jagged opening. Text bubbles say: "使えるのは!? 機械接続されてる昇降機だけ!" (What can be used!? Only the elevator connected by mechanical connection!) and "なら....." (Then.....).
- Bottom Center Panel:** A character stands next to a computer monitor displaying a map. Text says: "問題はこの運管理ソフト本物にも載ってるんですよ" (The problem is that this transportation management software is also loaded on the real thing).
- Bottom Right Panels:** Two smaller panels featuring a character with glasses talking on a phone. One panel says: "今日は降回路にシフトでも起こって電源が不安定になる..." (Today, even if there's a shift change on the circuit, the power becomes unstable...). The other panel shows two characters talking, with one saying: "それ引き金に黒澤君を死なせるんだぞ" (That's the trigger to make Kurosawa die, you know).

ダメですツツ
推力姿勢制御とも
不能!!

ちよつとこれ
どうなってるの!?
計測が……

はな
故障の
警告表示が
全部点いてる
こんなことって……ツ

反応しませんヨ

日向
はさ
す
ど
い
ゆ
こ
と
だ
杉田
田

使えるのは?

機械接続されてる
昇降舵だけ!

間題は
この海給管嘴ソフト
水物にも
熱つてるんですよ

なら……

それ引き金に
異常を起すん
だから

今日以降
船体にはシクでる船つて
電気が不安定になる……

なまじりとは
ソフトの入れ替えて
直りますかね

ALVO

ARCABOUÇO DO APRENDIZADO SUPERVISIONADO

DADOS DE TREINAMENTO

ENTRADA



RÓTULO



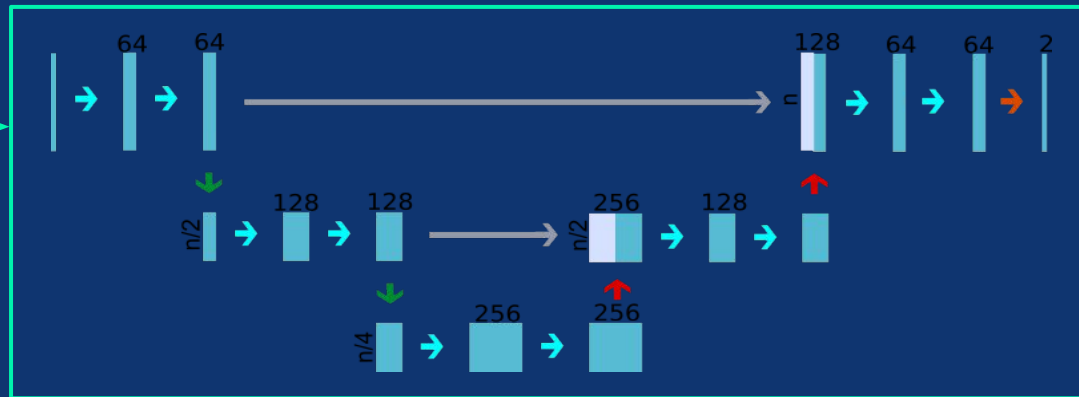
OTIMIZAÇÃO

FUNÇÃO DE PERDA



MODELO: U-NET

ENTRADA



PREDIÇÃO



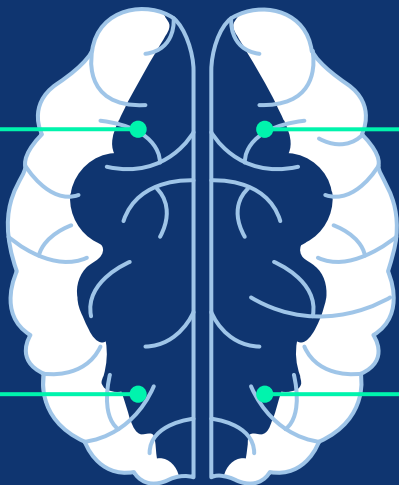
DESAFIOS

CAPACIDADE DA REDE

Overfit vs underfit;

VARIABILIDADE DE ESCRITAS

Os estilos dos caracteres podem variar muito em um mesmo mangá;



QUANTIDADE DE HIPERPARÂMETROS

Grande quantidade de hiperparâmetros para serem testados;

DESBALANCEAMENTO DE CLASSES

Aproximadamente 90% dos pixels são negativos.

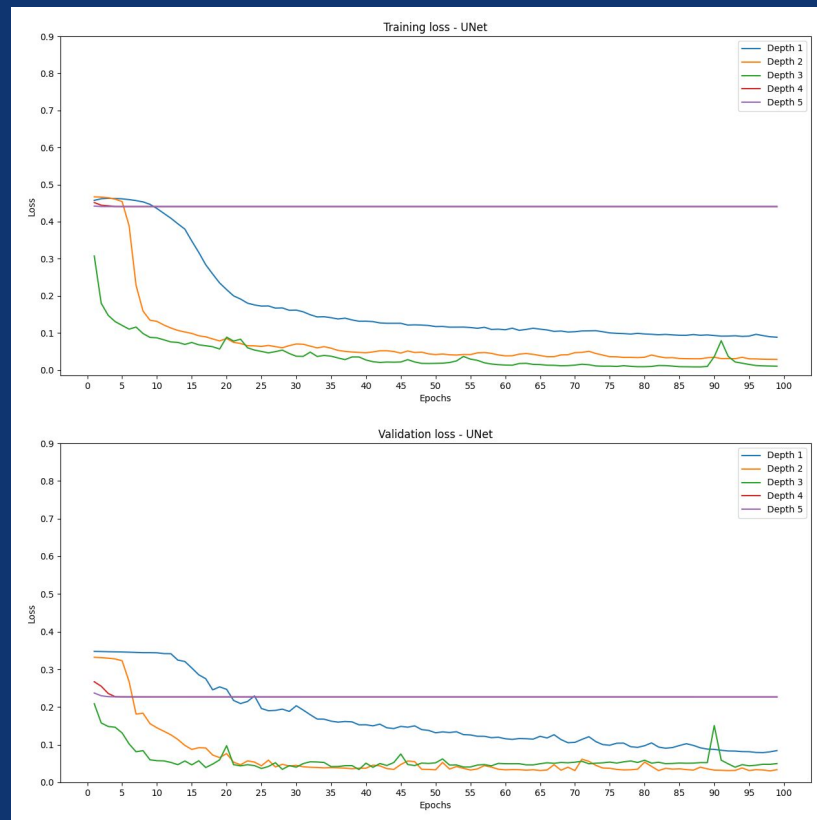
OBJETIVOS

- Treinar os modelos em contextos com poucos exemplos (<10);
- Ajustar hiperparâmetros para diminuir falsos positivos e falsos negativos;
- Investigar os impactos de cada hiperparâmetro.



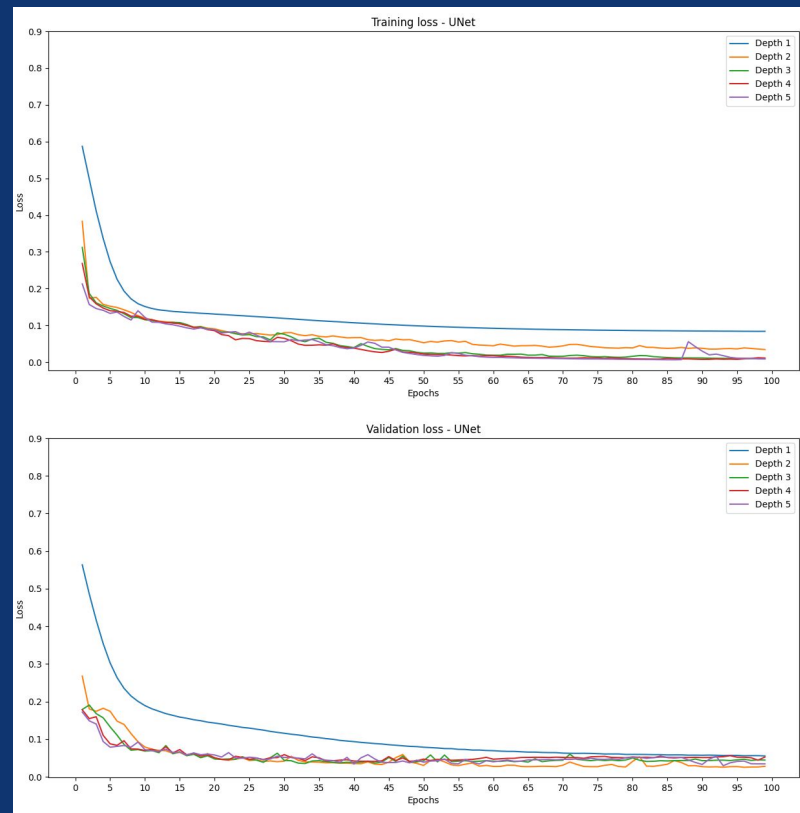
EXPERIMENTO 1: CAPACIDADE

	ACURÁCIA	PRECISÃO	RECALL
PROF. 1	0,988	0,433	0,193
PROF. 2	0,994	0,735	0,670
PROF. 3	0,993	0,753	0,578
PROF. 4	0,989	-	0,000
PROF. 5	0,989	-	0,000



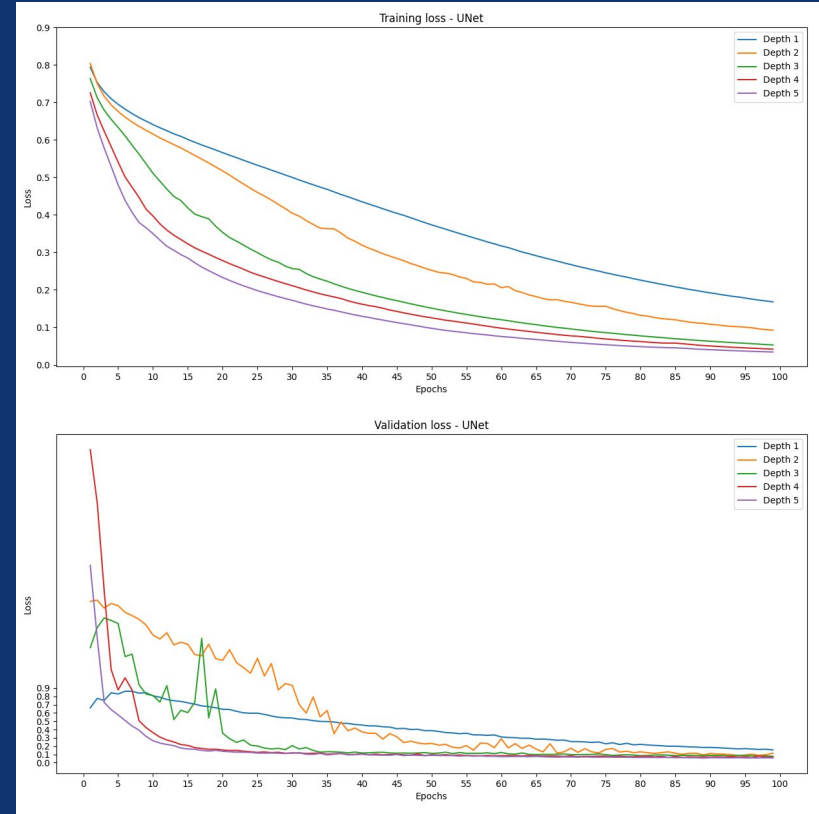
EXPERIMENTO 2: NORM. DE ENTRADA

	ACURÁCIA	PRECISÃO	RECALL
PROF. 1	0,989	0,484	0,020
PROF. 2	0,993	0,826	0,516
PROF. 3	0,990	0,932	0,090
PROF. 4	0,995	0,834	0,689
PROF. 5	0,997	0,941	0,763



EXPERIMENTO 3: NORM. DE *BATCH*

	ACURÁCIA	PRECISÃO	RECALL
PROF. 1	0,988	0,456	0,475
PROF. 2	0,996	0,843	0,747
PROF. 3	0,997	0,864	0,857
PROF. 4	0,998	0,917	0,869
PROF. 5	0,997	0,912	0,797



EXPERIMENTO 4: FUNÇÕES DE PERDA

	ACURÁCIA	PRECISÃO	<i>RECALL</i>
GDL	0,997	0,880	0,823
WCE	0,991	0,886	0,215
CE	0,990	0,932	0,090
FL	0,994	0,778	0,635

EXPERIMENTO 5: AVALIAÇÃO CRUZADA

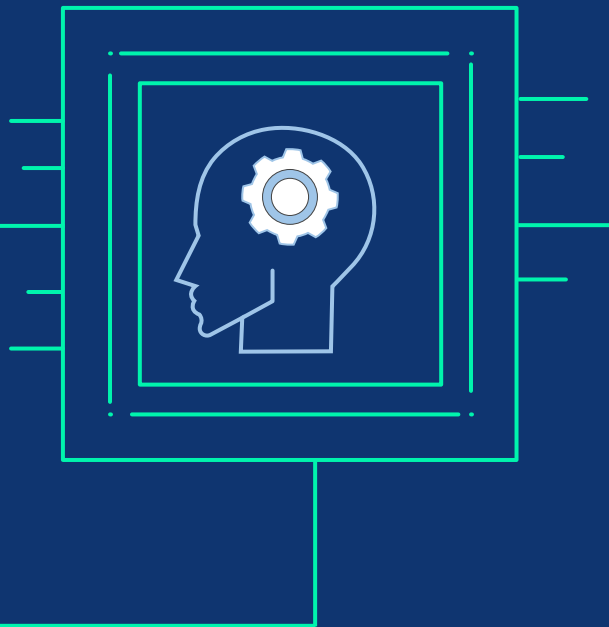
TREINO	TESTE	ACURÁCIA	PRECISÃO	RECALL
A	A	0,996	0,857	0,813
A	B	0,998	0,735	0,821
A	C	0,996	0,677	0,956
A	D	0,995	0,700	0,815

B	A	0,992	0,772	0,577
B	B	0,998	0,839	0,853
B	C	0,997	0,814	0,883
B	D	0,994	0,738	0,660

TREINO	TESTE	ACURÁCIA	PRECISÃO	RECALL
C	A	0,992	0,848	0,513
C	B	0,998	0,867	0,702
C	C	0,998	0,890	0,943
C	D	0,995	0,805	0,593

D	A	0,994	0,867	0,639
D	B	0,998	0,799	0,676
D	C	0,996	0,745	0,810
D	D	0,996	0,844	0,786

CONCLUSÕES



- Não existe estudo parecido com este;
- 10 páginas são suficientes;
- U-Net de profundidade 3 generaliza bem neste problema;
- Normalização de entrada traz benefícios e não prejudica tanto o tempo decorrido;
- A função de perda GDL contribui para uma melhor performance;
- Uma rede treinada em um título funciona bem em outro título;

OBRIGADO!

Pedro H. B. de Almeida
Estudante

Nina S. T. Hirata
Orientadora

Este projeto foi financiado pela Fundação de Amparo à
Pesquisa do Estado de São Paulo (FAPESP) sob o processo
2020/02891-3.

CREDITS: This presentation template was created by
Slidesgo, including icons by **Flaticon**, and infographics &
images by **Freepik**

