Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence (IJCAI-24)

Revisiting Neural Networks for Continual Learning: An Architectural Perspective

Aojun Lu[1] , **Tao Feng**[2] , **Hangjie Yuan**[3] , **Xiaotian Song**[1] and **Yanan Sun**[1*]
[1]Sichuan University
[2]Tsinghua University
[3]Zhejiang University
aojunlu@stu.scu.edu.cn, fengtao.hi@gmail.com, hj.yuan@zju.edu.cn
songxt@stu.scu.edu.cn, ysun@scu.edu.cn

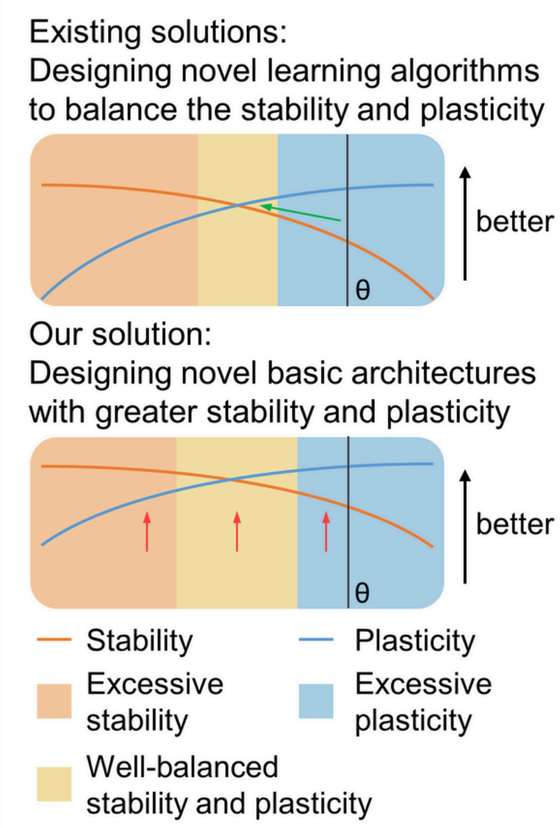
https://arxiv.org/pdf/2404.14829

SIN5006 - INTELIGÊNCIA COMPUTACIONAL
ENTREGA PARCIAL

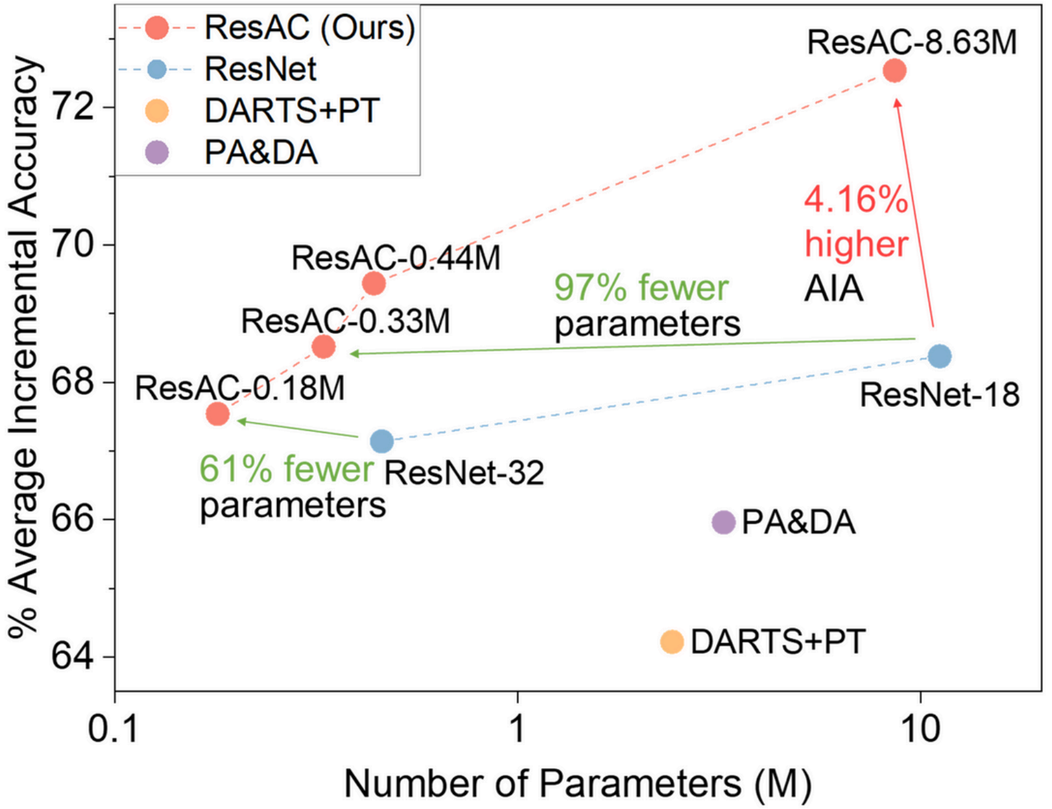
NILTON TADASHI ENTA - 12730911
VITOR CAMARGO PINHEIRO - 12693156

ArchCraft | Definição do problema

Apesar dos avanços em **Continous Learning (CL)**, o autor ressalta que há **menos atenção** dada à análise do papel do **design da arquitetura de rede** na contribuição da área



Motivation of ArchCraft



CL Performance vs. Number of Parameters

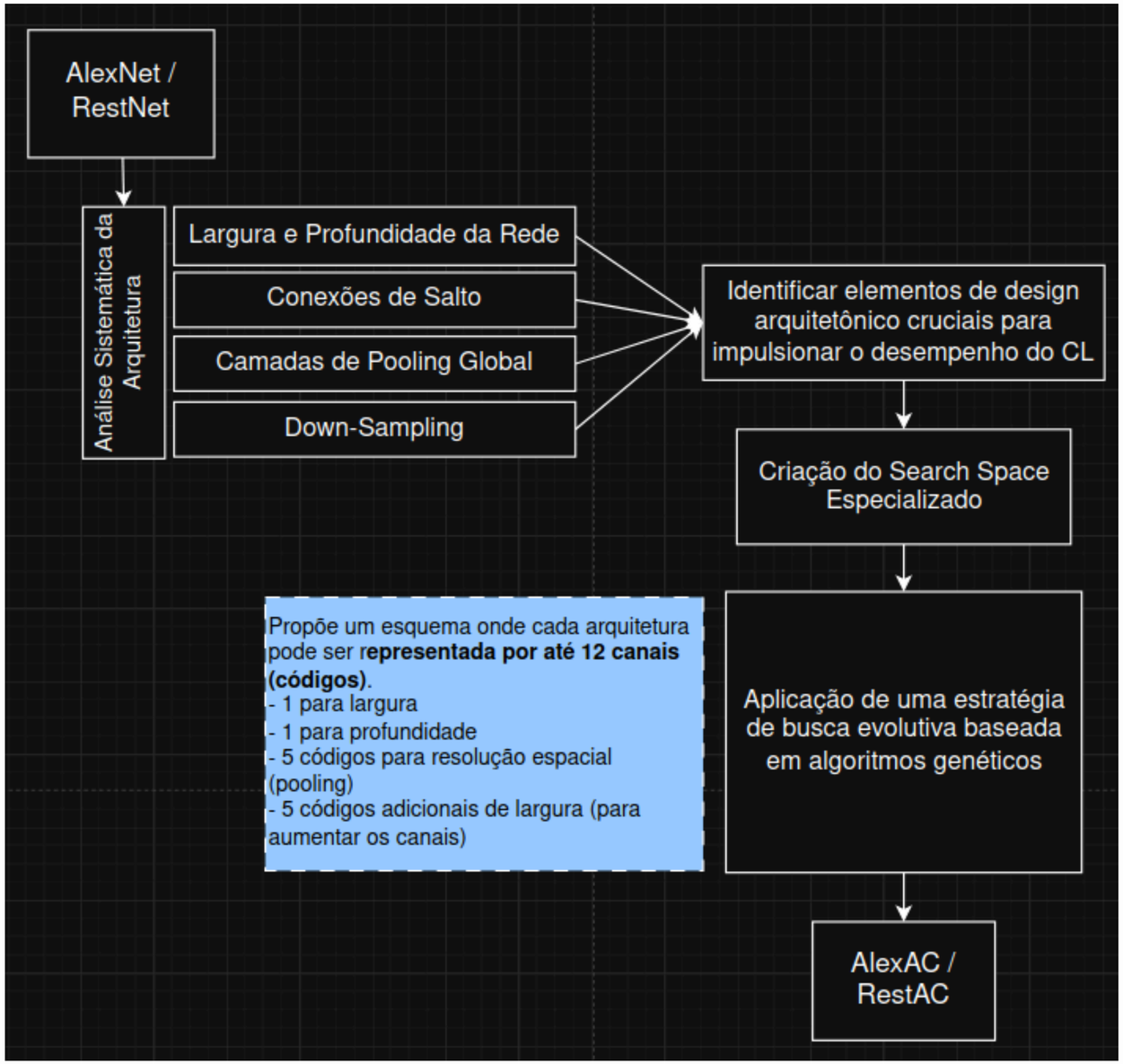
ArchCraft | Métodos (Visão Geral)

ArchCraft é um novo método de Neural Architecture Search (NAS).
Focado em Aprendizado Contínuo (CL) para preencher a lacuna entre o **design da arquitetura de rede e o CL.**

I. **Search Space**: Espaço de busca personalizado obtido a partir de **busca empírica** [1].

II. **Friendly Search Strategy:** Uma estratégia baseada em algoritmos genéticos para explorar o Search Space [2].

ArchCraft | Métodos (Visão Geral)



ArchCraft | Prometido vs Obtido

Prometido	Obtido
• Executar o projeto sob a mesma base de dados utilizada pelo autor	**EXECUTADO**
• Reproduzir o experimento do autor, seguindo os passos descritos no artigo	**PARCIAL:** Não foi reproduzido a comparação do ArchCarft com outros métodos para a tarefa de classificação

ArchCraft | Dificuldades

- Estudo e entendimento de conceitos e técnicas relacionadas à Redes Neurais
- Entendimento da AlexNet e ResNet
- Entendimento dos benchmarks
- Entendimento do código

ArchCraft | Desenho dos Experimentos

Conjunto de Dados

- CIFAR100
- Incremento de 5 classes (de 100) para CL

Medida de Avaliação

- Acurácia Incremental Média (AIA)
- Catastrophic Forgetting
- Benchmark

Seleção de Hiperparametros

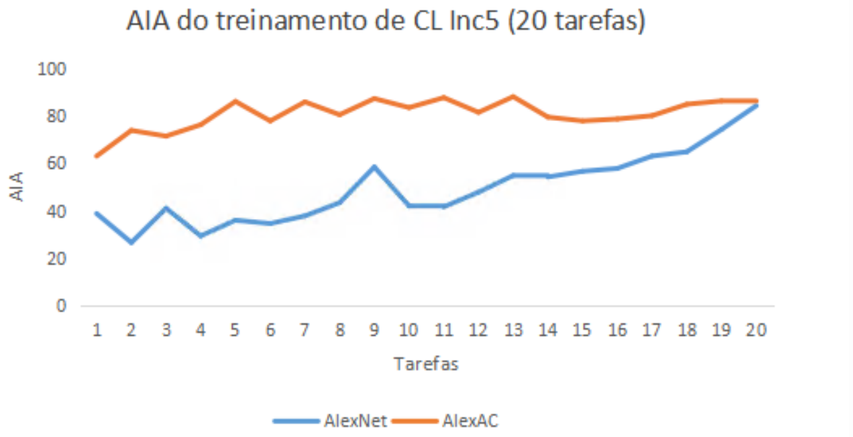
- **SGD**
  - lr: 0.01,
  - momentum: 0.9,
  - dampening: 0,
  - weight_decay: 0.0002,
  - nesterov: False,
  - maximize: False,
  - foreach: None,
  - differentiable: False

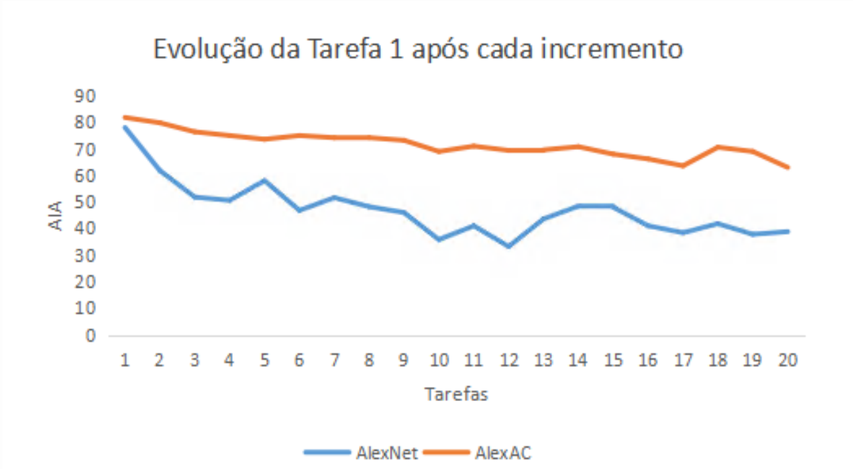
ArchCraft | Discussão dos Resultados

Medida	AlexNet	AlexAC	Resultado	
Acurácia Final	49,35	80,76	**31,41**	63,65 %
Cat. Forgetting	34,86	5,10	**-29,77**	-85,39 %
Params	6,71	6,28	**-0,43**	-6,43 %

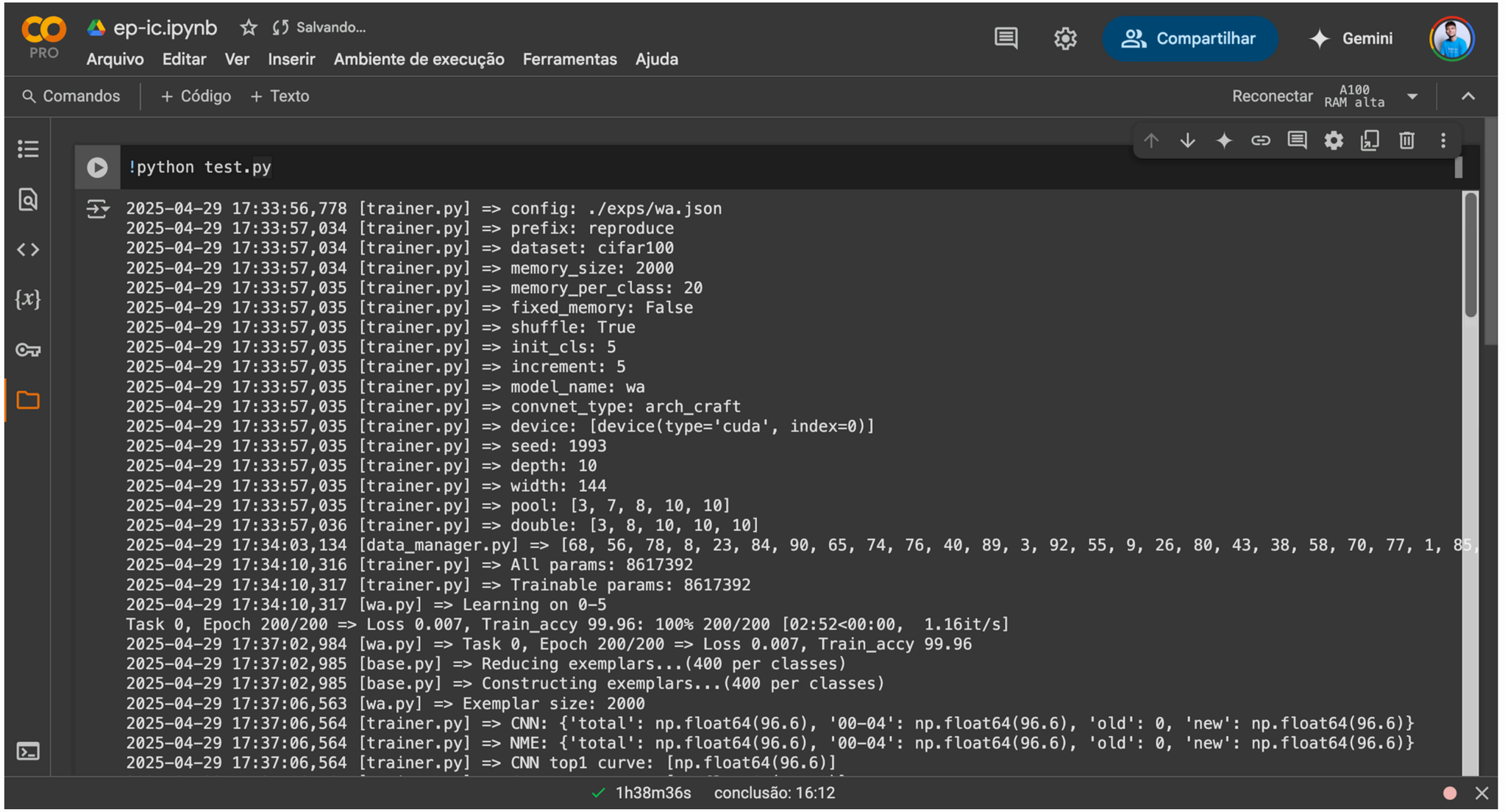
Network	#P (M)	Method	C100-inc5	C100-inc10
AlexNet	6.71	Upper Bound	81.86	73.07
		SGD	53.78	56.92
		SI	70.3	62.9
		HAT	71.8	62.8
		SPG	75.9	67.7
		WSN	76.9	69.3
AlexAC-A	6.28↓ 6%	Upper Bound	83.59	75.09
		SGD	**82.38** (+5.48)	**73.91** (+4.61)
AlexAC-B	0.92↓ 86%	Upper Bound	83.58	74.72
		SGD	79.32 (+2.42)	70.87 (+1.57)

Table 3: The last accuracy of the AlexAC and AlexNet in *Task IL*. '#P' represents the number of parameters of the network used. **Bolded** indicates best performance. Underline indicates second best.

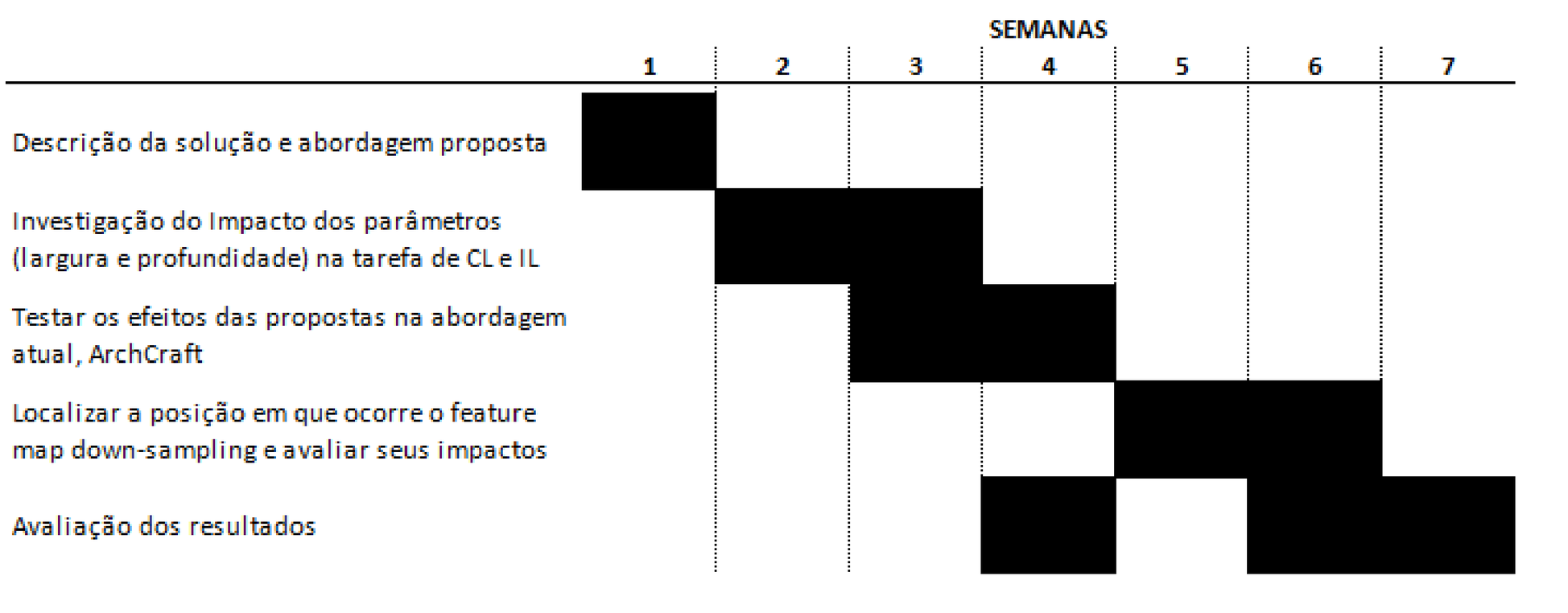




ArchCraft | Discussão dos Resultados



ArchCraft | Entrega Final e Cronograma



*Delimitar as análises para AlexNet

ArchCraft | Referências

LU, Aojun; FENG, Tao; YUAN, Hangjie; SONG, Xiaotian; SUN, Yanan. Revisiting neural networks for continual learning: an architectural perspective. arXiv:2404.14829, 2024. Disponível em: https://arxiv.org/abs/2404.14829. Acesso em: 30 abr. 2025.

ANEXOS

PSEUDOCODIGO

DEFINE VARIÁVEIS LOCAIS

incremento <- 50

individuo_id <- 0

code <- [profundidade, largura, pool[1, ..., 5], double[1, ..., 5]]

chosen_network <- 'arch_craft'

Algoritmo TrainModel:

grad_clip <- 10

epoch <- 2

learning_rate <- 0.01

Carrega os dados CIFAR-100 divididos em tarefas incrementais

Se rede == 'arch_craft': importa e instancia arquitetura Net personalizada

Se rede == 'alexnet': importa e instancia versão AlexNet

Mostra o número total de parâmetros do modelo

Inicializa o treinamento com otimizador SGD (Gradiente Descendente Estocástico)

Imprime a função de perda e o otimizador

Inicializa matrizes para armazenar acurácias, perdas, e medidas por tarefa

for tarefa **in** classes:

Treina o modelo

Salva os resultados

for conjunto **in** range(tarefas já processadas):

Testa a conjunto

Salva os resultados

Calcula a acurácia média após a tarefa

Calcula forgetting <- sum(acurácia máxima - t[x])/ len(t)

Imprime matriz de acurácia

Calcula e imprime métrica média ap (AIA)

Calcula e imprime acurácia

Retorna a acurácia final

ArchCraft | Métodos (Visão Geral)

ArchCraft é um novo método de Neural Architecture Search (NAS).
Focado em Aprendizado Contínuo (CL) para preencher a lacuna entre o **design da arquitetura de rede e o CL.**

I. **Search Space**: Espaço de busca personalizado obtido a partir de **busca empírica** [1].

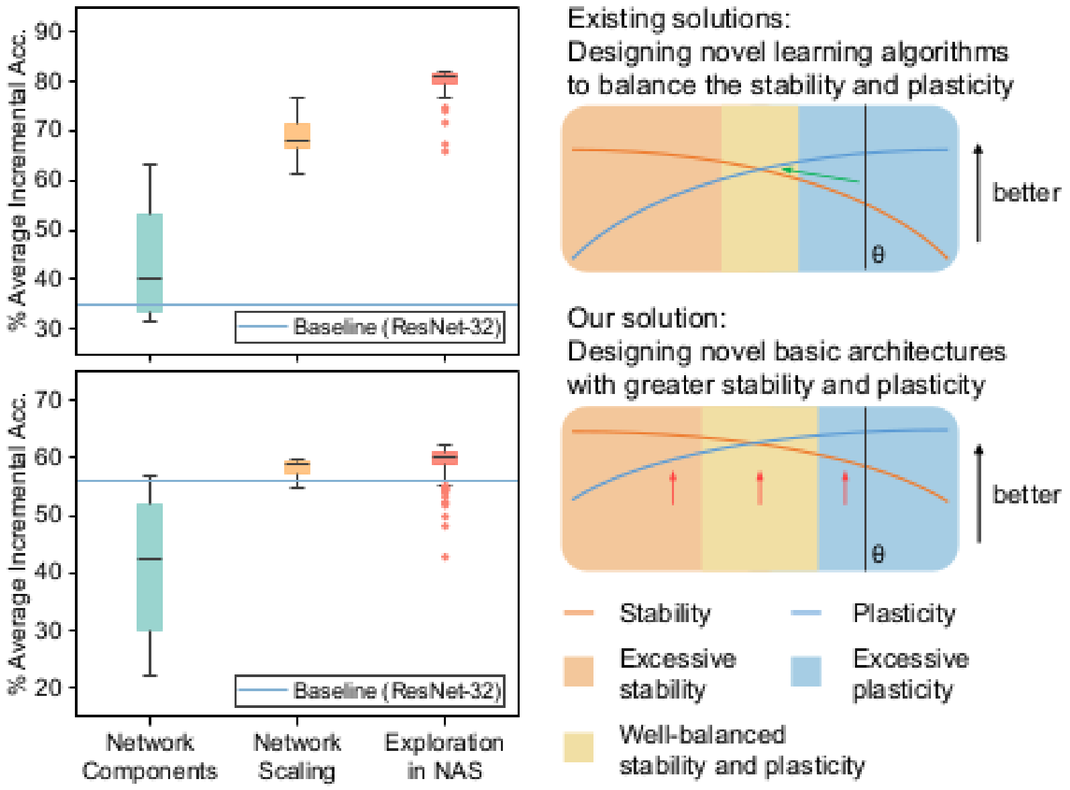
II. **Friendly Search Strategy:** Uma estratégia baseada em algoritmos genéticos para explorar o Search Space [2].

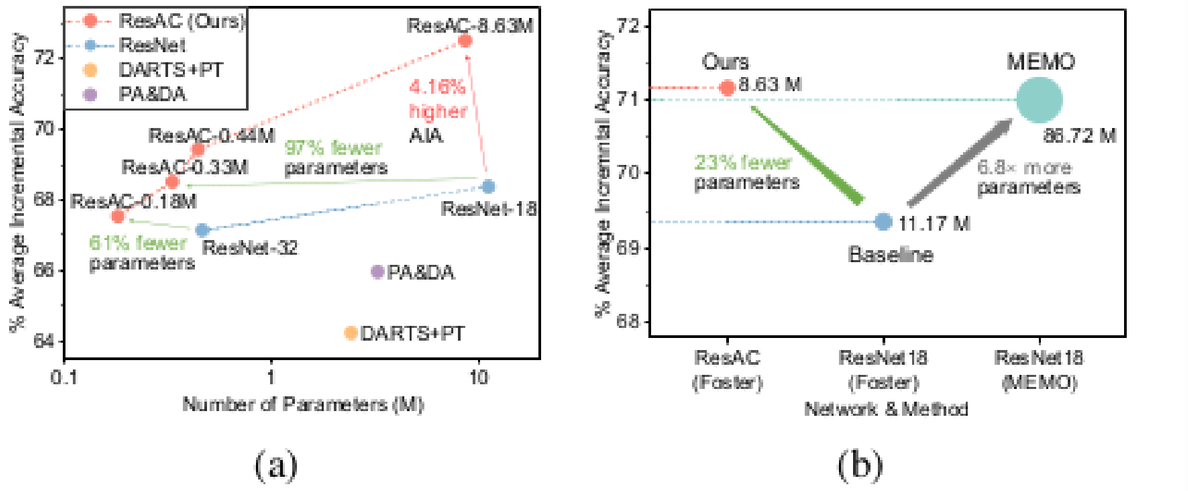
[1] O ArchCraft ele é projetado com base em **observações empíricas sobre seus componentes arquitetônicos**.
O experimento consistiu com uma ResNet. Partindo de configurações consideradas ótimas para essas tarefas (Incremental Learning (IL)): Max Pooling, Skip connection e Global Avg. Pooling

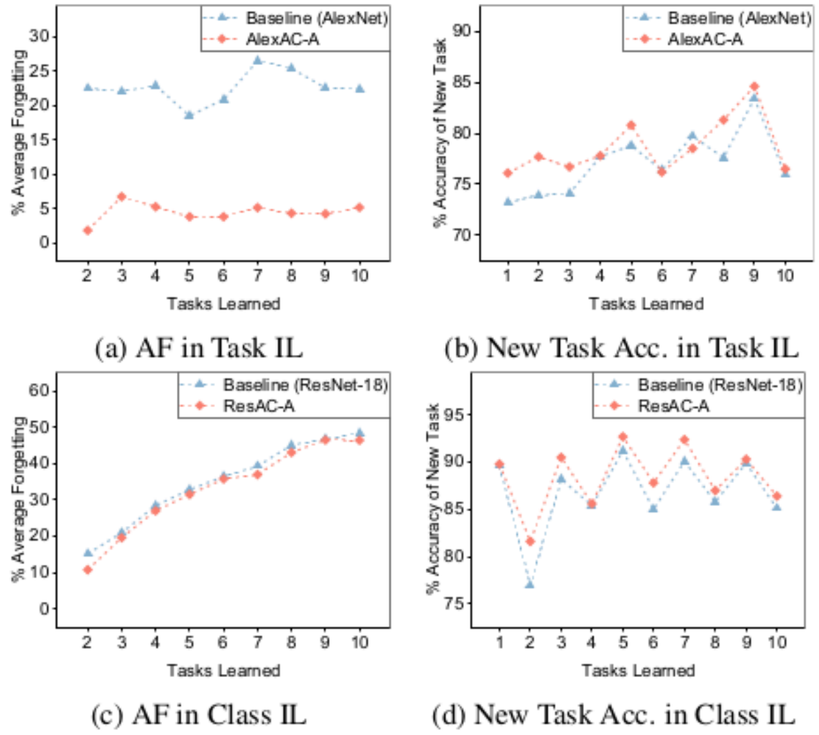
---

[2] **Inicialização:** Uma população de arquiteturas é gerada aleatoriamente
Avaliação da performance: durante as iterações evolutivas os indivíduos são avaliados por Acurácia Incremental Média (AIA).
Geração de novos indivíduos: novas arquiteturas são geradas a partir de dois indivíduos, escolhidos aleatoriamente, chamados de filho. O filho é resultado da mutação dessas duas arquiteturas.
Nova população: É gerada a partir dos filhos com melhor aptidão.

Anexos | Resultados







Anexos | Resultados

Method	Network	#P (M)	C100-inc5		C100-inc10		I100-inc5		I100-inc10		Max Improvement	
			LA	AIA	LA	AIA	LA	AIA	LA	AIA	LA	AIA
Replay	ResNet-32	0.46	39.10	58.17	40.02	58.21	-	-	-	-	-	-
	ResAC-B	0.44(↓ 4%)	**40.45**	**59.67**	**42.79**	**59.99**	-	-	-	-	+2.77	+1.78
	ResNet-18	11.17	40.04	58.80	43.23	60.42	36.30	57.30	41.00	59.21	-	-
	ResAC-A	8.63↓ 23%	**42.99**	**62.52**	**46.62**	**63.36**	**36.78**	**57.40**	**42.44**	**60.07**	+3.39	+3.72
iCaRL	ResNet-32	0.46	46.67	63.47	48.80	64.18	-	-	-	-	-	-
	ResAC-B	0.44↓ 4%	**47.94**	**64.17**	**50.11**	**64.42**	-	-	-	-	+1.31	+0.70
	ResNet-18	11.17	47.32	64.13	52.77	66.04	44.10	62.36	50.98	67.11	-	-
	ResAC-A	8.63↓ 23%	**52.6**	**68.71**	**55.52**	**69.62**	**45.12**	**63.98**	**52.46**	**68.42**	+5.28	+4.58
WA	ResNet-32	0.46	46.95	62.93	53.35	66.61	-	-	-	-	-	-
	ResAC-B	0.44↓ 4%	**51.31**	**66.39**	**54.89**	**67.73**	-	-	-	-	+4.36	+3.46
	ResNet-18	11.17	45.11	62.06	56.59	68.89	46.06	62.96	55.04	68.60	-	-
	ResAC-A	8.63↓ 23%	**53.23**	**69.19**	**59.79**	**71.40**	**49.94**	**67.20**	**58.86**	**71.56**	+8.12	+7.13
Foster	ResNet-32	0.46	47.78	62.36	54.36	67.14	-	-	-	-	-	-
	ResAC-B	0.44↓ 4%	**53.50**	**67.34**	**58.17**	**69.44**	-	-	-	-	+5.72	+4.98
	ResNet-18	11.17	49.03	61.97	55.98	68.38	53.26	65.20	60.58	69.36	-	-
	ResAC-A	8.63↓ 23%	**57.22**	**69.99**	**61.44**	**72.54**	**54.32**	**66.41**	**61.94**	**71.16**	+8.19	+8.02

Table 5: The CL performance of ArchCraft in *Class IL*. '#P' represents the number of parameters of the network used.