

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA
DEPARTAMENTO DE COMPUTAÇÃO

Sebastião Venâncio Guimarães Neto

**Aprendizado ativo profundo com comitê de
modelos**

São Carlos-SP

2025

Sebastião Venâncio Guimarães Neto

Aprendizado ativo profundo com comitê de modelos

Monografia apresentada ao Departamento de Computação da Universidade Federal de São Carlos como requisito parcial para obtenção do grau de Bacharel em Engenharia de Computação.

Orientação e Coorientação Profa. Dra. Priscila Tiemi Maeda Saito
Prof. Dr. Willian Paraguassu Amorim

São Carlos-SP

2025

Neto, Sebastião Venâncio Guimarães

Aprendizado ativo profundo com comitê de modelos /
Sebastião Venâncio Guimarães Neto -- 2025.
72f.

TCC (Graduação) - Universidade Federal de São Carlos,
campus São Carlos, São Carlos
Orientador (a): Priscila Tiemi Maeda Saito
Banca Examinadora: Alexandre Luis Magalhaes Levada,
Marcela Xavier Ribeiro, Willian Paraguassu Amorim
Bibliografia

1. Aprendizado ativo. 2. Aprendizado profundo. I. Neto,
Sebastião Venâncio Guimarães. II. Título.

Ficha catalográfica desenvolvida pela Secretaria Geral de Informática
(SIn)

DADOS FORNECIDOS PELO AUTOR

Bibliotecário responsável: Arildo Martins - CRB/8 7180

*À minha família, pelo amor e apoio em cada etapa desta jornada.
Aos meus professores e colegas de curso, pela troca de conhecimentos e incentivo constante.
E especialmente à minha orientadora, Profa. Dra. Priscila T. M. Saito, e ao meu
coorientador, Prof. Dr. Willian P. Amorim, pela paciência, sugestões valiosas e confiança
no meu trabalho.*

Agradecimentos

Agradeço, primeiramente, à minha orientadora, Profa. Dra. Priscila Tiemi Maeda Saito, e ao meu coorientador, Prof. Dr. Willian Paraguassu Amorim, pelo direcionamento, paciência e valiosas contribuições em todas as etapas deste trabalho.

À Universidade Federal de São Carlos e ao Departamento de Computação, pelo ambiente acadêmico de excelência e pelas oportunidades de aprendizado e pesquisa.

À minha família, pelo amor incondicional, compreensão e incentivo para que eu pudesse dedicar-me aos estudos com tranquilidade.

Aos amigos e colegas de curso, pelas discussões produtivas, apoio moral e companheirismo ao longo desta jornada.

Por fim, agradeço a todos que, direta ou indiretamente, contribuíram para a realização deste trabalho.

“Não aprendemos com a experiência, mas sim refletindo sobre ela.”
(John Dewey)

Resumo

A anotação manual de grandes volumes de dados permanece um dos principais gargalos em tarefas de visão computacional para treinamento de modelos de aprendizado profundo. Neste trabalho, propomos e avaliamos um comitê de estratégias de aprendizado ativo que combina critérios de seleção de amostras baseados em incerteza (Least Confidence, Margin Sampling e Entropy) e diversidade (k-means). Para validação experimental, utilizamos três bases com diferentes níveis de complexidade, MNIST, FashionMNIST e Parasitos, e adotamos métricas como acurácia, cobertura de classes conhecidas e percentual de amostras corrigidas. Os resultados indicam que o comitê atinge acurácias comparáveis às das melhores estratégias individuais, mas identifica todas as classes de maneira mais rápida, reduzindo o esforço total de anotação. Conclui-se que comitês em aprendizado ativo profundo promovem um equilíbrio eficaz entre exploração e diversidade, além de abrirem caminho para investigações futuras envolvendo bases mais desafiadoras, arquiteturas heterogêneas e cenários de adaptação de domínio.

Palavras-chave: Aprendizado Ativo; Comitê de Estratégias; Amostragem por Incerteza; Amostragem por Diversidade

Abstract

Manual annotation of large-scale datasets remains one of the main bottlenecks in computer vision tasks for training deep learning models. In this work, we propose and evaluate a committee of active learning strategies that combines uncertainty-based sample selection criteria (Least Confidence, Margin Sampling, and Entropy) with a diversity-based approach (k-means). For experimental validation, we employ three datasets with different levels of complexity, MNIST, FashionMNIST, and Parasitos, and adopt metrics such as accuracy, known-class coverage, and percentage of corrected samples. The results indicate that the committee achieves accuracies comparable to the best individual strategies while identifying all classes more rapidly, thus reducing the overall annotation effort. We conclude that committee-based deep active learning provides an effective balance between exploration and diversity, and further opens avenues for future investigations involving more challenging datasets, heterogeneous architectures, and domain adaptation scenarios.

Keywords: Active Learning; Strategy Committee; Uncertainty Sampling; Diversity Sampling

Lista de ilustrações

Figura 1 – Pipeline da metodologia proposta de comitês de estratégias de aprendizado ativo	20
Figura 2 – Exemplos de imagens de cada classe do conjunto de dados MNIST . . .	21
Figura 3 – Exemplos de imagens de cada classe do conjunto de dados FashionMNIST: (a) T-shirt/top, (b) Trouser, (c) Pullover, (d) Dress, (e) Coat, (f) Sandal, (g) Shirt, (h) Sneaker, (i) Bag, (j) Ankle boot	22
Figura 4 – Exemplos de imagens de cada classe do conjunto de dados Parasitos: (a) A. lumbricoides, (b) Ancilostomideo, (c) B. hominis, (d) E. coli, (e) E. histolytica, (f) E. nana, (g) E. vermicularis, (h) Eggs_Impurities, (i) Giardia, (j) H. diminuta, (k) H. nana, (l) I. butschlii, (m) Protozoans_Impurities, (n) S. mansoni, (o) S. stercoralis, (p) S. stercoralis_Impurities, (q) Taenia, (r) T. trichiura	23
Figura 5 – Arquitetura da rede neural utilizada nos experimentos, composta por duas camadas convolucionais, camada densa intermediária e camada de saída.	25
Figura 6 – Resultados obtidos por cada estratégia de aprendizado ativo utilizando o conjunto MNIST. (a) acurácias. (b) quantidade de classes conhecidas e (c) porcentagem de amostras corrigidas	30
Figura 7 – Resultados obtidos por cada estratégia de aprendizado ativo utilizando o conjunto FashionMNIST. (a) acurácias. (b) quantidade de classes conhecidas e (c) porcentagem de amostras corrigidas	33
Figura 8 – Resultados obtidos por cada estratégia de aprendizado ativo utilizando o conjunto Parasitos. (a) acurácias. (b) quantidade de classes conhecidas e (c) porcentagem de amostras corrigidas	35
Figura 9 – Resultados de acurácias obtidos pelo uso individual e pela combinação da estratégia de aprendizado ativo baseada em diversidade (KS) com as demais estratégias baseadas em incerteza: (a) MS, (b) LC e (c) EN, respectivamente, utilizando o conjunto de dados MNIST.	38
Figura 10 – Gráficos de radar sintetizando o desempenho das estratégias individuais e de suas combinações em termos de acurácia no teste, porcentagem de erros, número de classes conhecidas, tempo de seleção e tempo de treino+teste, utilizando o conjunto de dados MNIST. Cada subfigura compara um par de estratégias de aprendizado ativo e suas variantes em comitê: (a) KS–ES, (b) KS–MS, (c) LC–ES, (d) LC–KS, (e) LC–MS e (f) MS–ES. Todas as métricas foram normalizadas no intervalo [0, 1] para fins de comparação relativa.	39

Figura 11 – Resultados de acurárias obtidos pelo uso individual e pela combinação da estratégia de aprendizado ativo baseada em diversidade (KS) com as demais estratégias baseadas em incerteza: (a) MS, (b) LC, (c) EN, respectivamente, utilizando o conjunto de dados FashionMNIST.	43
Figura 12 – Gráficos de radar sintetizando o desempenho das estratégias individuais e de suas combinações em termos de acurácia no teste, porcentagem de erros, número de classes conhecidas, tempo de seleção e tempo de treino+teste, utilizando o conjunto de dados FashionMNIST. Cada subfigura compara um par de estratégias de aprendizado ativo e suas variantes em comitê: (a) KS–ES, (b) KS–MS, (c) LC–ES, (d) LC–KS, (e) LC–MS e (f) MS–ES. Todas as métricas foram normalizadas no intervalo $[0, 1]$ para fins de comparação relativa.	44
Figura 13 – Resultados de acurárias obtidos pelo uso individual e pela combinação da estratégia de aprendizado ativo baseada em diversidade (KS) com as demais estratégias baseadas em incerteza: (a) MS, (b) LC, (c) EN, respectivamente, utilizando o conjunto de dados Parasitos.	45
Figura 14 – Gráficos de radar sintetizando o desempenho das estratégias individuais e de suas combinações em termos de acurácia no teste, porcentagem de erros, número de classes conhecidas, tempo de seleção e tempo de treino+teste, utilizando o conjunto de dados Parasitos. Cada subfigura compara um par de estratégias de aprendizado ativo e suas variantes em comitê: (a) KS–ES, (b) KS–MS, (c) LC–ES, (d) LC–KS, (e) LC–MS e (f) MS–ES. Todas as métricas foram normalizadas no intervalo $[0, 1]$ para fins de comparação relativa.	46
Figura 15 – Resultados de acurárias obtidos pelo uso individual e pela combinação da estratégia de aprendizado ativo baseada em incerteza (LC) com as demais estratégias baseadas em incerteza: (a) ES, (b) KS, (c) MS, respectivamente, utilizando o conjunto de dados MNIST	52
Figura 16 – Resultados de acurária obtidos pela combinação da estratégia de aprendizado ativo baseada em entropia (ES) com as demais estratégias individuais: (a) MS e (b) KS, respectivamente, utilizando o conjunto de dados MNIST.	53
Figura 17 – Resultados de acurária obtidos pela combinação da estratégia de aprendizado ativo baseada em margem (MS) com a estratégia baseada em diversidade (KS), utilizando o conjunto de dados MNIST.	53
Figura 18 – Correções por iterações obtidas pela combinação da estratégia de aprendizado ativo baseada em incerteza (LC) com as estratégias individuais: (a) MS, (b) KS e (c) ES, respectivamente, utilizando o conjunto de dados MNIST.	54

Figura 19 – Correções por iterações obtidas pela combinação da estratégia de aprendizado ativo baseada em entropia (ES) com as estratégias individuais: (a) MS e (b) KS, respectivamente, utilizando o conjunto de dados MNIST.	55
Figura 20 – Correção por iterações obtida pela combinação da estratégia de aprendizado ativo baseada em margem (MS) com a estratégia KS, utilizando o conjunto de dados MNIST.	55
Figura 21 – Classes conhecidas ao longo das iterações utilizando a estratégia de aprendizado ativo baseada em incerteza (LC) combinada com: (a) KS, (b) MS e (c) ES no conjunto MNIST.	56
Figura 22 – Classes conhecidas ao longo das iterações utilizando a estratégia baseada em entropia (ES) combinada com: (a) KS e (b) MS no conjunto MNIST.	57
Figura 23 – Classes conhecidas ao longo das iterações utilizando a estratégia baseada em margem (MS) combinada com KS no conjunto MNIST.	57
Figura 24 – Acurácia obtida ao executar cada estratégia até atingir 150 iterações ou a acurácia considerando todo o conjunto de dados mais três rounds para o conjunto FashionMNIST.	58
Figura 25 – Resultados de acurácia obtidos pela combinação da estratégia de aprendizado ativo baseada em incerteza (LC) com as estratégias individuais: (a) ES, (b) KS e (c) MS, respectivamente, utilizando o conjunto de dados FashionMNIST.	59
Figura 26 – Resultados de acurácia obtidos pela combinação da estratégia de aprendizado ativo baseada em entropia (ES) com as estratégias individuais: (a) MS e (b) KS, respectivamente, utilizando o conjunto de dados FashionMNIST.	60
Figura 27 – Imagem do ensemble obtido pela combinação da estratégia de aprendizado ativo baseada em margem (MS) com KS, utilizando o conjunto de dados FashionMNIST.	60
Figura 28 – Correções por iterações obtidas pela combinação da estratégia de aprendizado ativo baseada em incerteza (LC) com as estratégias individuais: (a) MS, (b) KS e (c) ES, respectivamente, utilizando o conjunto de dados FashionMNIST.	61
Figura 29 – Correções por iterações obtidas pela combinação da estratégia de aprendizado ativo baseada em entropia (ES) com as estratégias individuais: (a) MS e (b) KS, respectivamente, utilizando o conjunto de dados FashionMNIST.	62
Figura 30 – Correção por iterações obtida pela combinação da estratégia de aprendizado ativo baseada em margem (MS) com KS, utilizando o conjunto de dados FashionMNIST.	62

Figura 31 – Classes conhecidas ao longo das iterações utilizando a estratégia de aprendizado ativo baseada em incerteza (LC) combinada com: (a) KS, (b) MS e (c) ES, respectivamente, no conjunto de dados FashionMNIST.	63
Figura 32 – Classes conhecidas ao longo das iterações utilizando a estratégia de aprendizado ativo baseada em entropia (ES) combinada com: (a) KS e (b) MS, respectivamente, no conjunto de dados FashionMNIST.	64
Figura 33 – Classes conhecidas ao longo das iterações utilizando a estratégia de aprendizado ativo baseada em margem (MS) combinada com KS, no conjunto de dados FashionMNIST.	64
Figura 34 – Acurácia obtida ao executar cada estratégia até atingir 150 iterações ou a acurácia considerando todo o conjunto de dados mais três rounds para o conjunto de Parasitos.	65
Figura 35 – Resultados de acurácia obtidos pela combinação da estratégia de aprendizado ativo baseada em incerteza (LC) com as estratégias individuais: (a) ES, (b) KS e (c) MS, respectivamente, utilizando o conjunto de dados Parasitos.	66
Figura 36 – Resultados de acurácia obtidos pela combinação da estratégia de aprendizado ativo baseada em entropia (ES) com as estratégias individuais: (a) MS e (b) KS, respectivamente, utilizando o conjunto de dados Parasitos.	67
Figura 37 – Resultados de acurácia obtidos pela combinação da estratégia de aprendizado ativo baseada em margem (MS) com KS, utilizando o conjunto de dados Parasitos.	67
Figura 38 – Correções por iterações obtidas pela combinação da estratégia de aprendizado ativo baseada em incerteza (LC) com as estratégias individuais: (a) MS, (b) KS e (c) ES, respectivamente, utilizando o conjunto de dados Parasitos.	68
Figura 39 – Correções por iterações obtidas pela combinação da estratégia de aprendizado ativo baseada em entropia (ES) com as estratégias individuais: (a) MS e (b) KS, respectivamente, utilizando o conjunto de dados Parasitos.	69
Figura 40 – Correção por iterações obtida pela combinação da estratégia de aprendizado ativo baseada em margem (MS) com KS, utilizando o conjunto de dados Parasitos.	69
Figura 41 – Classes conhecidas ao longo das iterações utilizando a estratégia de aprendizado ativo baseada em incerteza (LC) combinada com: (a) KS, (b) MS e (c) ES, respectivamente, no conjunto de dados Parasitos.	70
Figura 42 – Classes conhecidas ao longo das iterações utilizando a estratégia de aprendizado ativo baseada em entropia (ES) combinada com: (a) KS e (b) MS, respectivamente, no conjunto de dados Parasitos.	71

Figura 43 – Classes conhecidas ao longo das iterações utilizando a estratégia de aprendizado ativo baseada em margem (MS) combinada com KS, no conjunto de dados Parasitos.	71
---	----

Lista de tabelas

Tabela 1 – Resumo dos principais trabalhos relacionados em Active Learning. . . .	16
Tabela 2 – Número de amostras em cada classe do conjunto de dados MNIST . . .	21
Tabela 3 – Número de amostras em cada classe do conjunto de dados FashionMNIST	22
Tabela 4 – Número de amostras em cada classe do conjunto de dados Parasitos . .	22

Sumário

1	INTRODUÇÃO	14
2	REFERENCIAL TEÓRICO	16
3	METODOLOGIA	20
3.1	Conjuntos de Dados	20
3.2	Estratégias de Aprendizado Ativo	23
3.3	Arquitetura da Rede Neural	24
3.4	Protocolo Experimental	25
3.5	Métricas de Avaliação	26
4	RESULTADOS E DISCUSSÕES	28
4.1	Cenário 1 – Resultados Individuais das Estratégias de Aprendizado Ativo	29
4.1.1	Resultados em MNIST	29
4.1.2	Resultados em FashionMNIST	31
4.1.3	Resultados em Parasitos	32
4.2	Cenário 2 – Resultados das Combinações das Estratégias de Aprendizado Ativo	36
4.2.1	Resultados em MNIST	36
4.2.2	Resultados em FashionMNIST	39
4.2.3	Resultados em Parasitos	41
5	CONCLUSÃO	47
	REFERÊNCIAS	49
	APÊNDICES	51
	APÊNDICE A – MNIST	52
	APÊNDICE B – FASHIONMNIST	58
	APÊNDICE C – PARASITOS	65

1 Introdução

Nas últimas décadas, modelos de aprendizado profundo (*Deep Learning*) têm alcançado resultados de estado da arte em tarefas de visão computacional e processamento de linguagem natural. Entretanto, seu desempenho depende fortemente da disponibilidade de grandes volumes de dados rotulados, cuja anotação manual é demorada e onerosa. Esse fator constitui um dos principais obstáculos à aplicação prática de redes neurais profundas em contextos industriais e acadêmicos com recursos limitados para anotação.

O *Aprendizado Ativo* (*Active Learning* – AL) surge como uma alternativa promissora para mitigar esse problema ao selecionar iterativamente as amostras mais informativas para anotação (SETTLES, 2008). Nessa abordagem, o modelo identifica as instâncias que mais contribuem para seu refinamento, reduzindo significativamente o esforço de anotação sem comprometer a acurácia final. Com o avanço das redes neurais profundas, o AL evoluiu para o *Deep Active Learning* (DAL), que combina a expressividade de modelos complexos com critérios de seleção baseados em incerteza e diversidade (HOULSBY et al., 2011; GAL; GHAMRANI, 2017). Estratégias como *Least Confidence*, *Entropy Sampling* e *Margin Sampling* tornaram-se amplamente empregadas em aplicações de visão computacional e reconhecimento de padrões, apresentando resultados consistentes em cenários de dados limitados (LEWIS; GALE, 1994; ROY; MCCALLUM, 2001).

Apesar dos avanços, ainda há desafios relevantes em compreender quais estratégias de aprendizado ativo são mais eficazes para equilibrar custo de anotação e desempenho final do modelo, especialmente em tarefas com diferentes níveis de complexidade, desde conjuntos simples, como o MNIST (LECUN; CORTES; BURGESS, 2010), até bases mais complexas, como o FashionMNIST e o Parasitos. Um dos principais pontos de investigação é como combinar diferentes critérios de seleção de amostras para potencializar a eficiência do processo de anotação. Nesse sentido, estudos recentes indicam que abordagens baseadas em *ensembles* e comitês de estratégias podem aprimorar a diversidade das amostras e aumentar a robustez do processo de seleção (KRAWCZYK; CANO, 2019; POLYZOS; LU; GIANNAKIS, 2024; CHITTA et al., 2022; DU; ZHANG, 2005).

Diante desse contexto, o presente estudo tem como objetivo investigar e avaliar o potencial de técnicas de aprendizado ativo profundo adaptadas como comitês de estratégias, buscando reduzir o custo de anotação e otimizar o treinamento de redes neurais em cenários com dados limitados. Foram implementados e reproduzidos algoritmos clássicos de aprendizado ativo, como *Least Confidence*, *Entropy Sampling* e *Margin Sampling*, além de variantes com *dropout*, como *BALD* e *MarginSamplingDropout*, adaptadas para o contexto de redes neurais profundas. Também foram explorados métodos de diversificação, como

o *KMeansSampling*, combinados a critérios de incerteza, a fim de avaliar o potencial de comitês híbridos na seleção de amostras informativas. O desempenho das estratégias foi analisado por meio de métricas como acurácia, número de classes conhecidas e percentual de amostras corrigidas, utilizando três conjuntos de dados de diferentes níveis de complexidade: MNIST, FashionMNIST e Parasitos.

Este trabalho está organizado da seguinte forma. O [Capítulo 2](#) apresenta o referencial teórico, revisando os principais conceitos e abordagens de aprendizado ativo, com ênfase nas estratégias baseadas em incerteza, diversidade e suas variações em aprendizado profundo. O [Capítulo 3](#) descreve a metodologia adotada, detalhando os conjuntos de dados, a arquitetura da rede neural, o protocolo experimental e as métricas utilizadas para avaliação. Em seguida, o [Capítulo 4](#) apresenta os resultados e discussões obtidos nos experimentos realizados, analisando o desempenho das estratégias individuais e dos comitês propostos. Por fim, o [Capítulo 5](#) reúne as conclusões gerais e propõe direções para trabalhos futuros.

2 Referencial Teórico

Tabela 1 – Resumo dos principais trabalhos relacionados em Active Learning.

Paper	Year	Approach	Ensemble	Criterion	Dataset / Domain
(LEWIS; GALE, 1994)	1994	Traditional	–	Least Confidence	Text classification
(ROY; MCCALLUM, 2001)	2001	Traditional	–	Expected error reduction (Monte Carlo)	Text classification
(NGUYEN; SMEULDERS, 2004)	2004	Traditional	–	Diversity (pre-clustering)	Visual and textual data
(SETTLES, 2008)	2008	Traditional	–	Entropy, Margin Sampling	NLP / sequence labeling
(HOULSBY et al., 2011)	2011	Traditional	–	Mutual information (BALD)	Tabular and classification data
(DU; ZHANG, 2005)	2005	Traditional	✓	Average model ambiguity (boosting)	DOE
(WANG; JIN; DOHERTY, 2017)	2017	Traditional	✓	Query by Committee	PSO + surrogate models (Kriging)
(KRAWCZYK; CANO, 2019)	2019	Traditional	✓	Multi-Armed Bandit	Streams with concept drift
(POLYZOS; LU; GIANNAKIS, 2024)	2024	Traditional	✓	Weighted uncertainty (Gaussian)	Continuous regression (simulation)
(CHITTA et al., 2022)	2022	Deep AL	✓	Variance across checkpoints	ImageNet, CIFAR-10, CIFAR-100
(CHEN et al., 2023)	2023	Deep AL	✓	Uncertainty and complementarity (EAL-GAN)	20 benchmarks + 3 image datasets
(CARRETE et al., 2023)	2023	Deep AL	✓	Epistemic/aleatoric uncertainty	Physico-chemical simulations (DFT)

O Aprendizado Ativo (*Active Learning* – AL) tem se mostrado uma estratégia fundamental para reduzir custos de anotação em tarefas supervisionadas, principalmente em domínios onde o processo de rotulagem é caro, demorado ou exige conhecimento especializado. Diferente do aprendizado tradicional, que depende de grandes volumes de dados rotulados, o AL busca selecionar iterativamente as amostras mais informativas, maximizando o ganho de conhecimento do modelo a partir de um número limitado de exemplos anotados. Essa característica tem impulsionado sua adoção em diferentes áreas, como medicina, processamento de linguagem natural e visão computacional, onde os dados são abundantes, mas a rotulação é dispendiosa.

Historicamente, as primeiras estratégias de Aprendizado Ativo foram formuladas a partir de critérios de incerteza. (LEWIS; GALE, 1994) introduziram o método de *Least Confidence* (LC), que prioriza as amostras em que o modelo apresenta menor confiança na classe mais provável, partindo da hipótese de que instâncias com predições incertas fornecem informações valiosas sobre a fronteira de decisão. Posteriormente, (SETTLES, 2008) propuseram medidas mais robustas de incerteza, como a *Entropy Sampling*, que considera toda a distribuição de probabilidades de saída, e a *Margin Sampling*, que seleciona as amostras em que a diferença entre as duas classes mais prováveis é mínima, indicando

indecisão. Já (HOULSBY et al., 2011) ampliaram essa perspectiva ao propor o método *BALD* (*Bayesian Active Learning by Disagreement*), que utiliza a informação mútua entre os parâmetros e as predições do modelo para selecionar amostras que maximizem o ganho de informação, capturando tanto a incerteza dos dados quanto a incerteza epistemológica do próprio modelo. Esses métodos baseados em incerteza se tornaram os pilares do Aprendizado Ativo clássico e continuam sendo amplamente utilizados como referência e comparação em experimentos contemporâneos.

Com o avanço das técnicas de representação e o aumento da dimensionalidade dos dados, surgiram estratégias voltadas à diversidade. Em vez de selecionar apenas as amostras mais incertas, essas abordagens procuram garantir que o conjunto rotulado cubra de forma representativa o espaço de entrada. (NGUYEN; SMEULDERS, 2004) introduziram o uso de agrupamento prévio (*pre-clustering*) para esse fim, em que os dados não rotulados são divididos em grupos e os elementos mais próximos de cada centróide são escolhidos para anotação, assegurando que regiões distintas do espaço sejam exploradas. Abordagens como o *KMeans Sampling*, que segue o mesmo princípio de representatividade e cobertura, reforçam a importância da diversidade na escolha das amostras e reduzem a redundância entre exemplos selecionados. Essa transição da simples incerteza para a diversidade marcou um avanço importante, pois permitiu que o AL se tornasse mais eficiente em bases de dados complexas e heterogêneas.

Com o surgimento das redes neurais profundas, novos desafios surgiram em relação à estimativa de incerteza, já que modelos determinísticos tendem a ser mal calibrados. (GAL; GHARAMANI, 2017) demonstraram que o uso do *dropout* durante a inferência pode ser interpretado como uma aproximação bayesiana, permitindo quantificar incertezas epistemológicas. A partir dessa ideia, surgiram variações estocásticas das estratégias clássicas, como LC-Dropout, ES-Dropout e MS-Dropout, nas quais o modelo realiza múltiplas passagens com dropout ativado e calcula estatísticas (como média e variância) sobre as predições. Essa abordagem fornece uma medida mais confiável da incerteza, aproximando o comportamento de um comitê de modelos e conectando o Aprendizado Ativo tradicional ao paradigma de *ensembles bayesianos*.

Enquanto as abordagens clássicas se mostraram eficazes em diversos contextos, pesquisas recentes têm explorado a combinação de múltiplos modelos e critérios dentro de estruturas de *ensembles*, com o objetivo de aumentar a robustez e reduzir o viés da seleção ativa. Um dos primeiros trabalhos nesse sentido foi o de (DU; ZHANG, 2005), que aplicaram o conceito de comitê no contexto de *Design of Experiments* (DOE). Nesse estudo, a decisão de quais amostras rotular baseia-se na ambiguidade média entre os modelos que compõem o ensemble, de modo que as amostras com maior desacordo entre as predições são consideradas as mais informativas. Essa ideia abriu caminho para uma ampla gama de métodos de Aprendizado Ativo com múltiplos classificadores.

A evolução desse conceito pode ser observada em trabalhos como o de (KRAWCZYK; CANO, 2019), que propuseram o *Adaptive Ensemble Active Learning* para lidar com fluxos de dados sujeitos a *concept drift*. A técnica emprega o paradigma de *multi-armed bandit* para ajustar dinamicamente os pesos dos classificadores conforme seu desempenho recente, mantendo a adaptabilidade do sistema frente a mudanças na distribuição dos dados. Já no domínio do aprendizado profundo, (CHITTA et al., 2022) ampliaram o uso de ensembles ao propor o *Training Data Subset Search with Ensemble Active Learning*, em que centenas de pontos de verificação (*checkpoints*) intermediários de redes neurais são reutilizados como membros de um ensemble implícito. Essa abordagem permite escalar o Aprendizado Ativo para grandes bases como CIFAR-10, CIFAR-100 e ImageNet, mostrando que é possível obter subconjuntos de treinamento mais eficientes do que o conjunto completo, tanto em desempenho quanto em tempo de treinamento.

Seguindo essa mesma linha, (POLYZOS; LU; GIANNAKIS, 2024) introduziram o *Weighted Ensembles for Adaptive Active Learning*, um ensemble de modelos Gaussianos com pesos dinâmicos ajustados conforme o desempenho recente. O método reforça a importância de ponderar a contribuição de cada modelo e demonstra que a combinação adaptativa de diferentes estimadores de incerteza pode gerar ganhos substanciais em tarefas de regressão contínua. Também explorando o potencial dos comitês, (WANG; JIN; DOHERTY, 2017) aplicaram o Aprendizado Ativo em contextos de otimização custosa, integrando o critério de desacordo de comitê (*Query by Committee*) a algoritmos de *Particle Swarm Optimization* e modelos substitutos, reduzindo significativamente o número de avaliações necessárias de funções caras de otimizar.

Mais recentemente, o Aprendizado Ativo tem sido incorporado em arquiteturas de aprendizado profundo mais sofisticadas. (CHEN et al., 2023) propuseram o *EAL-GAN* (*Ensemble Active Learning Generative Adversarial Network*), que combina redes generativas condicionais e múltiplos discriminadores cooperativos. A arquitetura introduz uma função de perda colaborativa entre discriminadores e um mecanismo ativo de amostragem, reduzindo o custo de anotação para cerca de 6% dos dados originais e superando métodos anteriores em mais de vinte conjuntos de dados de detecção de anomalias. De forma complementar, (CARRETE et al., 2023) compararam *deep ensembles* e comitês clássicos no contexto de estimação de incerteza, demonstrando que ensembles profundos apresentam melhor calibração e correlação entre erro e incerteza, sendo particularmente adequados para ciclos iterativos de aprendizado ativo.

Essa trajetória evidencia uma clara evolução na literatura: o campo passou das estratégias pontuais baseadas em incerteza para abordagens mais abrangentes que combinam diversidade, adaptabilidade e aprendizado profundo. As contribuições mais recentes reforçam que a integração de múltiplos critérios e modelos, sejam eles bayesianos, gaussianos ou neurais, tende a equilibrar exploração e exploração durante o processo de seleção

ativa, aumentando a eficiência do aprendizado supervisionado com menos dados anotados. É nesse contexto que o presente trabalho se insere, buscando compreender e aprimorar o uso de ensembles e estratégias clássicas de Aprendizado Ativo em modelos de visão computacional, explorando a intersecção entre teoria e prática para reduzir o custo de anotação sem comprometer o desempenho dos modelos.

3 Metodologia

O presente projeto apresenta a proposta e a avaliação de um comitê de estratégias de aprendizado ativo, o qual combina critérios de seleção de amostras baseadas em diversidade e incerteza. São avaliadas três formas de combinação: sequencial, aplicando uma estratégia de aprendizado ativo seguida de outra; intersecção e união de estratégias. A Figura 1 apresenta o passo a passo da metodologia proposta.

A metodologia adotada neste trabalho foi elaborada de modo a permitir a análise sistemática e reproduzível do comportamento de diferentes estratégias de aprendizado ativo em redes neurais profundas. O experimento foi conduzido com base em um protocolo padronizado de amostragem incremental, implementado integralmente em Python e disponível publicamente no repositório do GitHub¹. Nesta seção, são descritos os conjuntos de dados utilizados, as estratégias de aprendizado ativo avaliadas, a arquitetura da rede neural empregada, o protocolo experimental adotado e as métricas de avaliação.

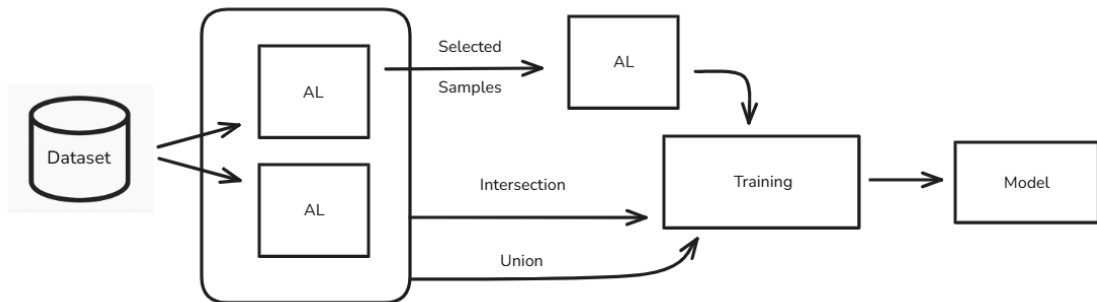


Figura 1 – Pipeline da metodologia proposta de comitês de estratégias de aprendizado ativo

3.1 Conjuntos de Dados

Foram utilizados três conjuntos de dados, os amplamente utilizados **MNIST**, **FashionMNIST**, (XIAO; RASUL; VOLLGRAF, 2017), e o **Parasitos**, composto por dados reais de microscopia, representando um cenário mais próximo de aplicações práticas.

O **MNIST** contém imagens de dígitos manuscritos com dimensões de 28×28 pixels em escala de cinza, sendo composto por 60.000 amostras de treinamento e 10.000 de teste, distribuídas em dez classes (0–9) (LECUN; CORTES; BURGES, 2010). A distribuição de amostras por classe é apresentada na Tabela 2, enquanto exemplos representativos podem ser vistos na Figura 2.

¹ <<https://github.com/TiaoVenancio/TCC-ActiveLearning>>

O **FashionMNIST** apresenta as mesmas dimensões e número de amostras, porém com imagens de peças de vestuário, sendo um benchmark mais desafiador por envolver maior variação intra-classe. A Tabela 3 resume a quantidade de instâncias por categoria, e a Figura 3 ilustra exemplos visuais de cada classe.

Já o **Parasitos** consiste em 41.332 imagens RGB de 200×200 pixels, divididas em 18 classes de parasitos microscópicos. Esse último conjunto foi incluído para avaliar o desempenho das estratégias em um cenário mais realista, com classes desbalanceadas e características visuais complexas. A Tabela 4 apresenta o número de imagens por classe, e a Figura 4 mostra amostras representativas de cada categoria.

Cada conjunto de dados passou por um processo de pré-processamento padronizado: conversão para *tensor* com valores normalizados entre 0 e 1, centralização estatística com média 0.1307 e desvio padrão 0.3081, e divisão entre conjuntos de treino e teste conforme as proporções originais, de acordo com Huang (2021).

Tabela 2 – Número de amostras em cada classe do conjunto de dados MNIST

Class	Training	Test
0	5923	980
1	6742	1135
2	5958	1032
3	6131	1010
4	5842	982
5	5421	892
6	5918	958
7	6265	1028
8	5851	974
9	5949	1009

Figura 2 – Exemplos de imagens de cada classe do conjunto de dados MNIST



Tabela 3 – Número de amostras em cada classe do conjunto de dados FashionMNIST

Class	Total Images
T-shirt/top	7000
Trouser	7000
Pullover	7000
Dress	7000
Coat	7000
Sandal	7000
Shirt	7000
Sneaker	7000
Bag	7000
Ankle boot	7000

Figura 3 – Exemplos de imagens de cada classe do conjunto de dados FashionMNIST: (a) T-shirt/top, (b) Trouser, (c) Pullover, (d) Dress, (e) Coat, (f) Sandal, (g) Shirt, (h) Sneaker, (i) Bag, (j) Ankle boot

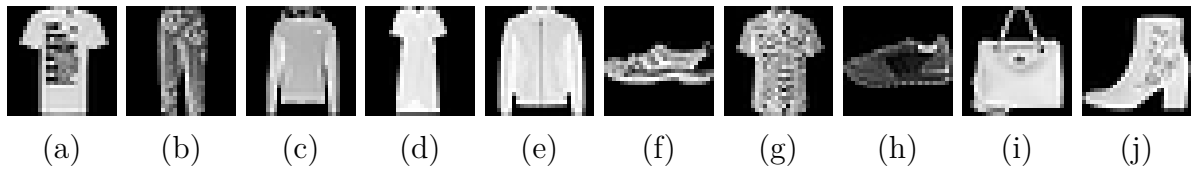
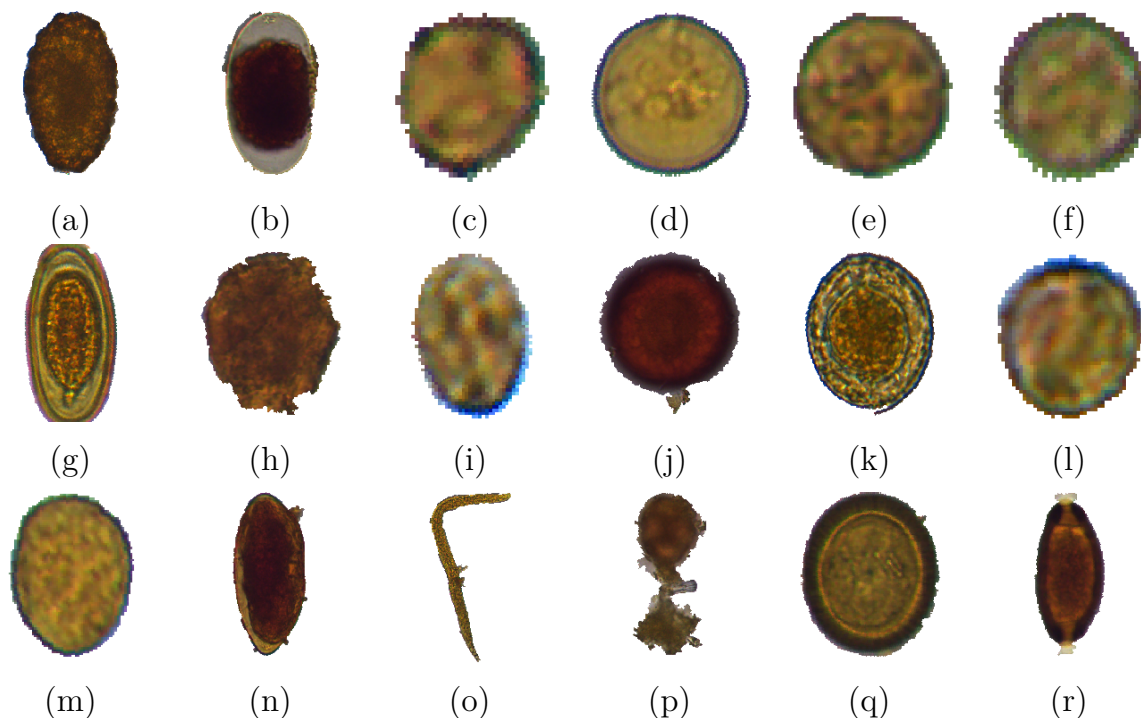


Tabela 4 – Número de amostras em cada classe do conjunto de dados Parasitos

Class	Total Images
A.lumbricoides	835
Ancilostomideo	286
B.hominis	309
E.coli	869
E.histolytica	659
E.nana	1783
E.vermicularis	103
Eggs_Impurities	9815
Giardia	1931
H.diminuta	83
H.nana	501
I.butschlii	3297
Protozoans_Impurities	28528
S.mansoni	254
S.stercoralis	247
S.stercoralis_Impurities	1351
T.trichiura	435
Taenia	379

Figura 4 – Exemplos de imagens de cada classe do conjunto de dados Parasitos: (a) *A. lumbricoides*, (b) *Ancilostomideo*, (c) *B. hominis*, (d) *E. coli*, (e) *E. histolytica*, (f) *E. nana*, (g) *E. vermicularis*, (h) *Eggs_Impurities*, (i) *Giardia*, (j) *H. diminuta*, (k) *H. nana*, (l) *I. butschlii*, (m) *Protozoans_Impurities*, (n) *S. mansoni*, (o) *S. stercoralis*, (p) *S. stercoralis_Impurities*, (q) *Taenia*, (r) *T. trichiura*



3.2 Estratégias de Aprendizado Ativo

Neste trabalho, foram avaliadas estratégias clássicas de *Active Learning* (AL), amplamente utilizadas na literatura para reduzir o custo de anotação e acelerar o processo de aprendizado. Essas abordagens podem ser organizadas em dois grandes eixos conceituais: métodos baseados em incerteza e métodos baseados em diversidade.

Os **métodos baseados em incerteza** selecionam amostras sobre as quais o modelo apresenta maior dúvida em suas predições. Entre eles, destaca-se o *Least Confidence* (LC) (LEWIS; GALE, 1994), que escolhe instâncias cuja maior probabilidade prevista pelo modelo é a mais baixa. O *Entropy Sampling* (ES) (SETTLES, 2008) utiliza a entropia da distribuição de probabilidades para quantificar essa incerteza global, enquanto o *Margin Sampling* (MS) considera a diferença entre as duas maiores probabilidades previstas, priorizando amostras com margens reduzidas. Já o método *BALD* (HOULSBY et al., 2011) emprega a informação mútua entre parâmetros e predições para identificar instâncias que tendem a gerar maior ganho informacional sobre o modelo.

Do ponto de vista operacional, a incerteza é calculada diretamente a partir das probabilidades previstas pelo modelo após a aplicação da função softmax. No método *Least Confidence*, a incerteza de uma amostra é associada à maior probabilidade prevista,

priorizando instâncias com baixa confiança máxima. No Margin Sampling, considera-se a diferença entre as duas classes mais prováveis, de modo que margens menores indicam maior ambiguidade na decisão do classificador. Já no Entropy Sampling, a incerteza é quantificada pela entropia da distribuição de probabilidades, calculada como a soma ponderada dos logaritmos das probabilidades previstas. Em todos os casos, as amostras com maiores níveis relativos de incerteza são selecionadas para anotação.

Complementando esse grupo, foram incluídas versões estocásticas das técnicas de incerteza, baseadas em *dropout* ativo durante a inferência (GAL; GHARAMANI, 2017). Nessas variantes (LC-Dropout, ES-Dropout e MS-Dropout), o modelo realiza múltiplas predições para cada amostra, produzindo distribuições empíricas que capturam a variabilidade das decisões e conferem maior robustez à estimativa de incerteza.

O segundo grupo é composto por **métodos baseados em diversidade**, cujo objetivo é selecionar instâncias que aumentem a cobertura do espaço de dados e reduzam redundâncias. Nesse contexto, foi utilizado o *KMeans Sampling* (KS) (NGUYEN; SMEULDERS, 2004), que realiza o agrupamento dos vetores de representação (*embeddings*) do conjunto não rotulado e seleciona, de cada grupo, o elemento mais próximo do centróide, assegurando a escolha de amostras representativas e variadas.

Essas estratégias, utilizadas individualmente ou combinadas, formam o conjunto de abordagens avaliadas neste trabalho para analisar o impacto da seleção ativa no desempenho de modelos convolucionais.

3.3 Arquitetura da Rede Neural

A rede neural utilizada segue uma arquitetura clássica de duas camadas convolucionais, similar à implementada em Huang (2021), conforme ilustrado na Figura 5. A primeira camada aplica 10 filtros de tamanho 5×5 , seguida de ativação ReLU e *max pooling* 2×2 . A segunda camada repete essa configuração com 20 filtros, incorporando um *Dropout2D* para mitigar o *overfitting*. O mapa de características resultante é então achatado em um vetor de 320 componentes e processado por uma camada totalmente conectada ($320 \rightarrow 50$), que gera um *embedding* de 50 dimensões, posteriormente utilizado por estratégias baseadas em diversidade como o *KMeans Sampling*. A camada final ($50 \rightarrow 10$), com ativação *softmax*, fornece a distribuição probabilística sobre as classes. O treinamento é conduzido com a função de perda *CrossEntropyLoss* e o otimizador Adam, empregando taxa de aprendizado de 10^{-3} .

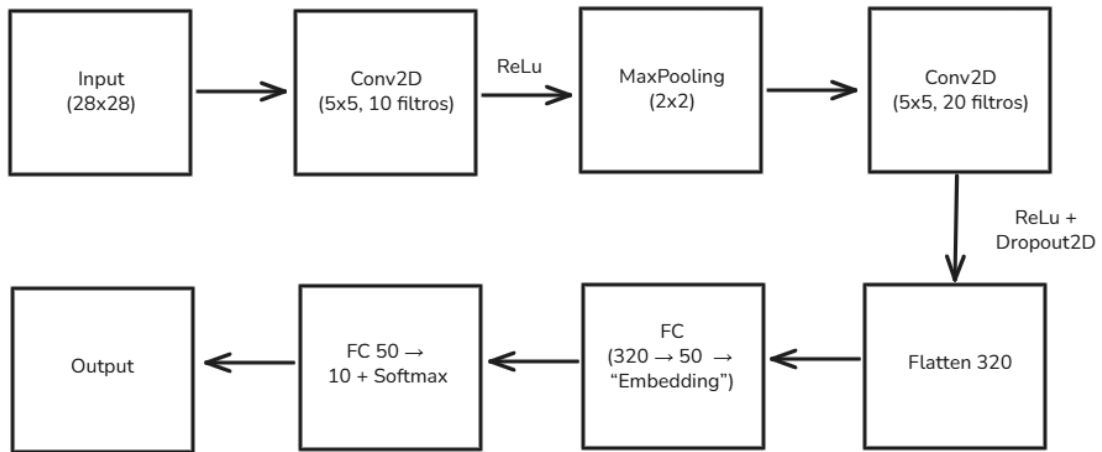


Figura 5 – Arquitetura da rede neural utilizada nos experimentos, composta por duas camadas convolucionais, camada densa intermediária e camada de saída.

3.4 Protocolo Experimental

O experimento foi conduzido em ambiente controlado, utilizando GPU NVIDIA RTX 3060 (12 GB) e imagem Docker pré-configurada com Ubuntu, Python, PyTorch e CUDA 12.4.1. Inicialmente, um conjunto reduzido de amostras rotuladas ($N_0 = 200$) foi selecionado aleatoriamente como *seed set*. O restante das instâncias compôs o *pool* não rotulado.

Algoritmo 1: Ciclo geral de Aprendizado Ativo

Input: Conjuntos rotulado L e não rotulado U ; modelo M ; número de iterações T ; tamanho da consulta n ; método de seleção \mathcal{S}

Output: Modelo treinado e métricas de desempenho

Round 0: Treinar o modelo M em L e registrar métricas iniciais;

for $t \leftarrow 1$ **to** T **do**

$J \leftarrow \mathcal{S}(M, U, n)$; // Seleciona amostras pelo método escolhido

Consultar o oráculo para obter os rótulos y_J ;

Atualizar $L \leftarrow L \cup \{(J, y_J)\}$ e $U \leftarrow U \setminus J$;

Re-treinar o modelo M usando todo o conjunto rotulado L ;

Avaliar M no conjunto de teste e registrar métricas da iteração;

end

Esse protocolo, ilustrado na Figura 1 e detalhado nos Algoritmos 1-4, foi repetido para cada conjunto de dados e para todas as estratégias (individuais e combinadas), garantindo comparabilidade entre execuções.

Algoritmo 2: Comitê de Seleção por Pré-filtro e Refinamento

Input: Modelo M ; pool não rotulado U ; tamanho da consulta n ; tamanho expandido m ; estratégias \mathcal{P} (pré-filtro) e \mathcal{R} (refinamento)

Output: Conjunto J com n amostras selecionadas

$I_P \leftarrow \mathcal{P}(M, U, m)$; // Pré-filtro gera um conjunto ampliado

Construir subconjunto U_P contendo apenas as amostras indexadas por I_P ;

$I_R \leftarrow \mathcal{R}(M, U_P, n)$; // Refinamento atua apenas no subconjunto

Mapear índices locais I_R para índices globais $J \subseteq I_P$;

return J

Algoritmo 3: Comitê de Seleção por Intersecção (Consenso)

Input: Modelo M ; pool não rotulado U ; tamanho da consulta n ; tamanho expandido m ; estratégias \mathcal{A} e \mathcal{B}

Output: Conjunto J com n amostras selecionadas

$I_A \leftarrow \mathcal{A}(M, U, m)$;

$I_B \leftarrow \mathcal{B}(M, U, m)$;

$I_\cap \leftarrow I_A \cap I_B$;

if $|I_\cap| \geq n$ **then**

$J \leftarrow$ primeiros n elementos de I_\cap ;

end

else

$C \leftarrow (I_A \cup I_B) \setminus I_\cap$;

 Completar $J \leftarrow I_\cap \cup$ primeiros $(n - |I_\cap|)$ elementos de C ;

end

return J

Algoritmo 4: Comitê de Seleção por União (Cobertura)

Input: Modelo M ; pool não rotulado U ; tamanho da consulta n ; tamanho expandido m ; estratégias \mathcal{A} e \mathcal{B}

Output: Conjunto J com n amostras selecionadas

$I_A \leftarrow \mathcal{A}(M, U, m)$;

$I_B \leftarrow \mathcal{B}(M, U, m)$;

$I_\cup \leftarrow I_A \cup I_B$;

if $|I_\cup| \geq n$ **then**

$J \leftarrow$ primeiros n elementos de I_\cup ;

else

$R \leftarrow U \setminus I_\cup$;

 Amostrar aleatoriamente de R até completar n elementos;

$J \leftarrow I_\cup \cup R$;

end

return J

3.5 Métricas de Avaliação

A avaliação das estratégias de aprendizado ativo foi conduzida com base em métricas selecionadas para capturar o desempenho preditivo do modelo, o comportamento

dinâmico do processo de seleção de amostras ao longo das iterações e o custo computacional. A primeira métrica considerada é a **acurácia no conjunto de teste**, utilizada para quantificar o desempenho do modelo após cada iteração de treinamento. Essa medida permite observar a evolução da capacidade de generalização à medida que novas amostras são progressivamente incorporadas ao conjunto rotulado.

Outra métrica considerada é o **número de classes conhecidas**, que indica quantas classes diferentes já apareceram no conjunto rotulado até cada iteração. Essa informação é particularmente relevante para cenários com desbalanceamento ou presença de classes raras, como no conjunto de dados Parasitos, permitindo analisar se determinada estratégia favorece uma descoberta mais rápida e ampla da diversidade do conjunto de dados.

A métrica referente à **porcentagem de amostras corrigidas** avalia a estabilidade das rotulações ao longo do processo. Essa métrica quantifica a proporção de instâncias cujo rótulo atribuído pelo modelo foi alterado em iterações posteriores, refletindo o grau de consistência das predições conforme o modelo se especializa. Uma taxa elevada de correção pode indicar incerteza persistente ou dificuldades específicas do conjunto de dados, enquanto valores mais baixos sinalizam maior estabilidade nas classificações. Além disso, essa métrica evidencia o impacto de cada estratégia de aprendizado ativo na capacidade de selecionar amostras realmente mais informativas para o aprendizado do modelo.

Por fim, as métricas relacionadas à mensuração de tempos computacionais (de seleção, treinamento e teste). Todas essas métricas, analisadas em conjunto, permitem uma caracterização abrangente do comportamento das estratégias de aprendizado ativo, contemplando não apenas o desempenho final, mas também a dinâmica de descoberta, estabilização e evolução do modelo ao longo das iterações.

4 Resultados e Discussões

Este capítulo apresenta e discute os resultados obtidos nos experimentos realizados com as estratégias de Aprendizado Ativo descritas no Capítulo 3. O objetivo central é analisar como cada abordagem, individualmente ou em comitê, influencia o desempenho do modelo ao longo das iterações, considerando tanto o ganho em acurácia quanto a eficiência no uso de amostras rotuladas.

Inicialmente, são apresentados os resultados referentes ao uso individual de cada estratégia de Aprendizado Ativo, aplicadas separadamente aos três conjuntos de dados considerados, **MNIST**, **FashionMNIST** e **Parasitos**, respectivamente. Em seguida, são discutidos os desempenhos obtidos pelos comitês formados pela combinação dessas estratégias.

Em todos os cenários, o modelo foi inicializado com um conjunto reduzido de amostras rotuladas e atualizado progressivamente à medida que novas instâncias são selecionadas pelas estratégias de Aprendizado Ativo.

A avaliação dos resultados baseia-se nas métricas previamente definidas, que permitem acompanhar a evolução do modelo sob diferentes perspectivas. A acurácia no conjunto de teste indica o desempenho ao final de cada iteração, enquanto o número de classes conhecidas mensura a velocidade com que o modelo passa a reconhecer novas categorias presentes nos dados. O custo de anotação acumulado expressa o total de instâncias rotuladas ao longo do processo, refletindo a eficiência da seleção ativa. A porcentagem de amostras corrigidas, medida a cada iteração, possibilita analisar a estabilidade das rotulações e o comportamento da estratégia em relação à consistência das seleções. Além disso, os tempos de seleção e de treinamento e teste também foram obtidos e permitem avaliar o trade-off entre desempenho preditivo e custo computacional.

Com base nessas observações, este capítulo busca identificar quais estratégias apresentam os melhores desempenhos em cada conjunto de dados, compreender de que forma a complexidade dos dados influencia o comportamento dos métodos e avaliar se combinações entre critérios distintos, como incerteza e diversidade, são capazes de superar abordagens isoladas. Também são investigadas quais técnicas fornecem o melhor equilíbrio entre economia de anotações e desempenho final do modelo.

Nas seções seguintes, são apresentados detalhadamente os resultados obtidos para cada conjunto de dados, descrevendo os padrões observados ao longo das iterações e destacando particularidades de cada estratégia. Em seguida, são analisados os métodos combinados e, ao final, é apresentada uma discussão integrada que sintetiza as principais conclusões empíricas do estudo.

4.1 Cenário 1 – Resultados Individuais das Estratégias de Aprendizado Ativo

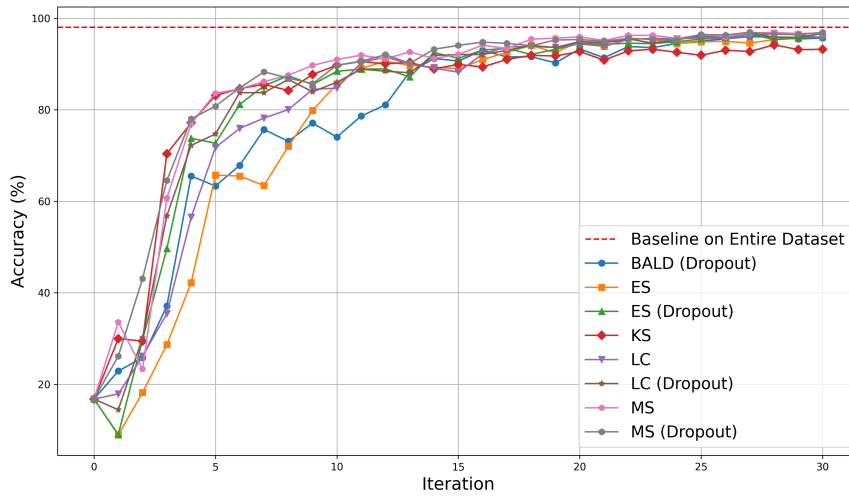
Nesta Seção são analisados os resultados obtidos por cada estratégia individual de *Active Learning* (AL) aplicada a cada conjunto de dados. A análise concentra-se no comportamento isolado de cada técnica em relação à acurácia, ao ritmo de descoberta de classes e à estabilidade das rotulações ao longo das iterações, permitindo observar como diferentes critérios de seleção influenciam o processo de aprendizado em cenários com distintos níveis de complexidade visual.

As Figuras 6–34 ilustram a evolução dessas métricas em cada conjunto de dados. Como referência, também é incluído o modelo totalmente supervisionado, treinado com 100% das amostras rotuladas, que alcança aproximadamente 98% de acurácia no MNIST, 91% no FashionMNIST e 94% no Parasitos. As estratégias de AL, em contraste, atingem níveis equivalentes de desempenho empregando cerca de 17,5% das instâncias anotadas, evidenciando a eficiência do processo de seleção ativa e o ganho expressivo em economia de esforço de anotação. As Subseções 4.1.1-4.1.3, apresentam a discussão detalhada dos resultados obtidos em cada conjunto de dados (MNIST, FashionMNIST e Parasitos, respectivamente), destacando regularidades, variações e particularidades observadas entre as estratégias avaliadas.

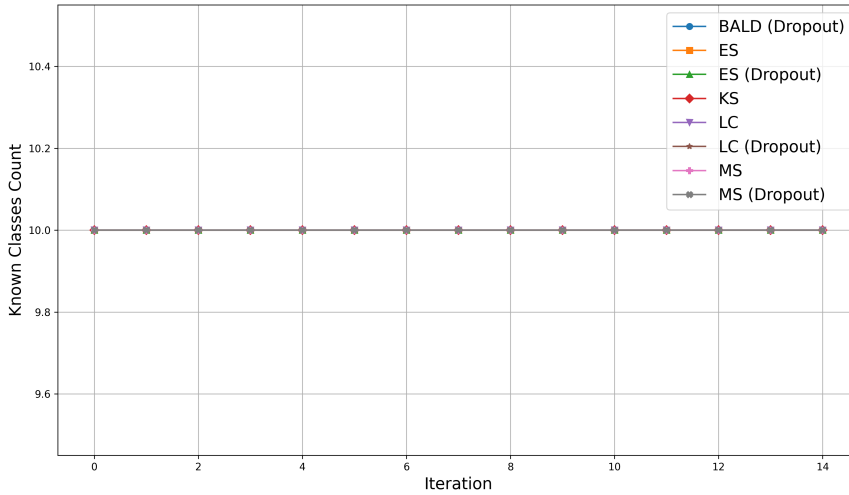
4.1.1 Resultados em MNIST

O gráfico da Figura 6(a) apresenta a evolução da acurácia ao longo das iterações para cada estratégia individual de AL aplicada ao conjunto de dados MNIST. De modo geral, todas as técnicas exibem crescimento acelerado nas primeiras iterações, convergindo para valores próximos ao desempenho do modelo totalmente supervisionado – i.e. em torno de 98%, considerando todas as 40000 amostras do conjunto – após cerca de 25 a 30 ciclos de seleção. Nesse ponto, o modelo treinado via AL, por exemplo, especificamente pela estratégia MS, utiliza apenas cerca de 13.00% a 15.50% do conjunto rotulado, equivalente a 6200 amostras, evidenciando o ganho expressivo em eficiência de anotação proporcionado pelo processo de seleção ativa.

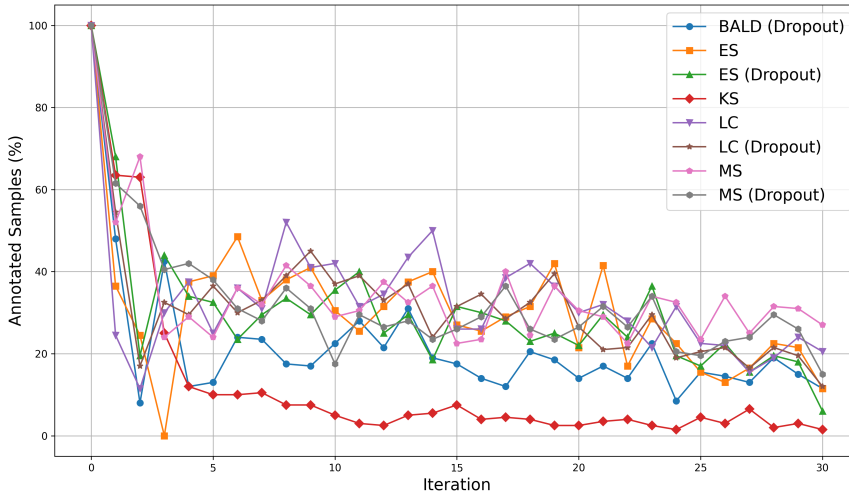
Dentre as estratégias de aprendizado ativo, as abordagens baseadas em diversidade e margem, especialmente a *KMeans Sampling* (KS) e a *Margin Sampling* (MS), demonstram crescimento inicial mais rápido, atingindo valores elevados de acurácia ainda nas primeiras iterações. Esse comportamento indica que, em cenários simples como o MNIST, a seleção de instâncias mais dispersas (no caso da KS) ou próximas da fronteira de decisão (como ocorre na MS) acelera o processo de aprendizado nos experimentos iniciais. Por outro lado, as estratégias baseadas em incerteza, como *Least Confidence* (LC) e *Entropy Sampling* (ES), exibem crescimento mais gradual nas iterações iniciais. Apesar do avanço inicial mais lento,



(a)



(b)



(c)

Figura 6 – Resultados obtidos por cada estratégia de aprendizado ativo utilizando o conjunto MNIST. (a) acurácias. (b) quantidade de classes conhecidas e (c) porcentagem de amostras corrigidas

essas técnicas alcançam acurácias finais superiores às obtidas pela KS, aproximando-se do desempenho máximo do modelo totalmente supervisionado, em torno de 98%. Entre as estratégias avaliadas, LC, ES e MS convergem para valores muito próximos desse limite, enquanto a KS se estabiliza ligeiramente abaixo. Esse comportamento indica que a seleção orientada por diversidade tende a ser mais vantajosa nas primeiras etapas, quando o modelo ainda necessita de exemplos amplamente representativos, mas mostra menor efetividade na fase final, momento em que critérios baseados em incerteza capturam melhor as regiões decisivas para refinamento do aprendizado.

A Figura 6(b) mostra que todas as estratégias identificam rapidamente as dez classes presentes no conjunto de dados desde a primeira iteração. Isso ocorre devido à distribuição equilibrada e da alta separabilidade entre as classes, permitindo que mesmo um conjunto inicial reduzido de amostras rotuladas já contenha representantes da maior parte das classes, independentemente do critério de seleção ativo empregado.

A análise da Figura 6(c), conforme esperado, revela um comportamento distinto daquele observado nos resultados de acurácias obtidas. Vale destacar que os valores das acurácias apresentadas referem-se aos desempenhos obtidos por cada modelo a cada iteração aplicados ao conjunto de teste, já os valores de porcentagem de amostras corrigidas referem-se às amostras selecionadas a cada iteração por cada estratégia de aprendizado ativo, i.e. ao conjunto de treinamento sendo incrementado a cada iteração. A estratégia KS apresenta a menor porcentagem de amostras corrigidas ao longo das iterações, refletindo sua tendência a selecionar exemplos mais representativos, mas menos ambíguos, característica que contribui para uma maior estabilidade preditiva. As estratégias baseadas em incerteza exibem um comportamento intermediário, com um volume moderado de correções distribuído ao longo do processo, coerente com sua ênfase em selecionar amostras nas quais o modelo demonstra dúvida. Nesse contexto, a estratégia MS destaca-se por apresentar o maior percentual de amostras corrigidas, especialmente nas iterações finais, evidenciando sua forte propensão a selecionar instâncias próximas da fronteira de decisão. Esse padrão indica que a estratégia prioriza exemplos mais ambíguos ou difíceis, que demandam ajustes sucessivos do modelo à medida que o processo de aprendizado evolui.

4.1.2 Resultados em FashionMNIST

Devido à maior sobreposição entre classes e às variações sutis de textura presentes nas imagens, o FashionMNIST configura-se como o cenário mais desafiador entre os conjuntos avaliados. O modelo totalmente supervisionado alcança cerca de 91% de acurácia quando treinado com todas as amostras do conjunto de dados (40000 instâncias). Em contraste, as estratégias de *Active Learning* convergem para níveis equivalentes de desempenho utilizando apenas uma fração do conjunto rotulado. Um exemplo ilustrativo é a estratégia MS, que atinge aproximadamente a mesma acurácia na iteração 99 empregando cerca de

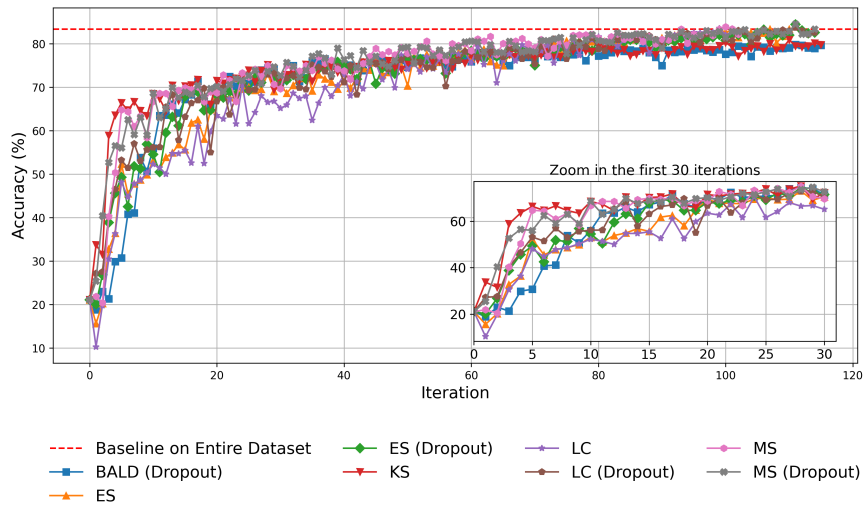
50.00% dos dados anotados (20000 amostras), evidenciando a eficiência do processo de seleção ativa mesmo em um domínio de maior complexidade visual (Figura 7(a)).

Comportamentos semelhantes aos observados no MNIST também foram verificados no conjunto FashionMNIST. As estratégias baseadas em diversidade e em fronteira de decisão, particularmente *KMeans Sampling* (KS) e *Margin Sampling* (MS), apresentaram as melhores trajetórias nas iterações iniciais, alcançando acurácias elevadas com maior rapidez e mantendo desempenho competitivo ao longo de todo o processo. Em contraste, as estratégias de incerteza pura, *Entropy Sampling* (ES) e *Least Confidence* (LC), exibiram progresso mais lento nas primeiras iterações. Embora a ES gradualmente se aproxime do desempenho das demais estratégias após um número maior de ciclos de seleção, a LC permanece sistematicamente abaixo das outras abordagens em praticamente todo o intervalo analisado. Esse comportamento sugere que, em cenários com maior sobreposição entre classes e variação intrínseca nas texturas, características marcantes do FashionMNIST, a incerteza isolada não constitui um critério suficientemente discriminativo para identificar amostras realmente informativas para o modelo.

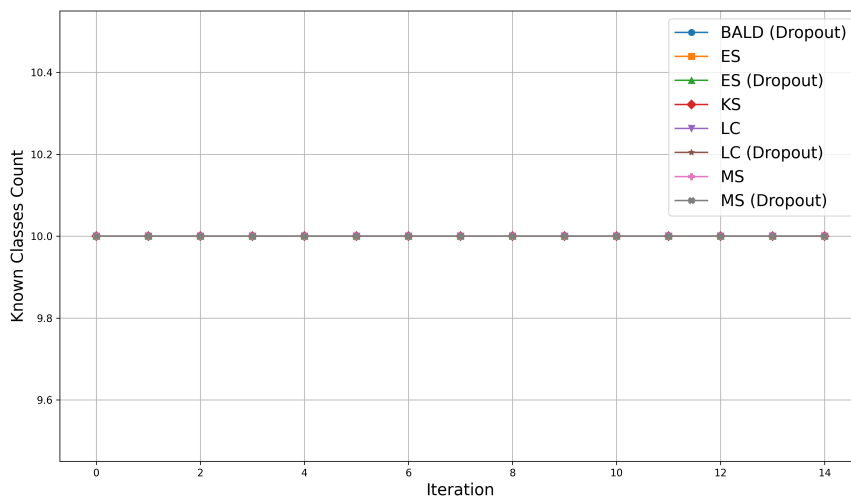
Com relação à quantidade de classes conhecidas (Figura 7(b)) e à quantidade de correções (Figura 7(c)) realizadas ao longo das iterações, também pode-se verificar, para o FashionMNIST, os mesmos comportamentos observados com o conjunto MNIST. Todas as estratégias apresentam cobertura total das classes desde a primeira iteração, dado que o FashionMNIST possui distribuição equilibrada entre categorias, facilitando a descoberta completa do conjunto de classes independentemente do critério de seleção inicial. Em termos de correções nos rótulos fornecidos pelo modelo, a estratégia KS apresenta o menor percentual de amostras corrigidas, mantendo uma trajetória estável e pouco oscilante. Já as estratégias ES, LC e MS exibem curvas de correção mais elevadas e com maior variabilidade, caracterizadas por picos e quedas ao longo do processo. Tal comportamento indica que essas estratégias tendem a selecionar instâncias mais ambíguas ou próximas da fronteira de decisão, o que gera ajustes frequentes nas predições do modelo. Apesar dessa instabilidade, tanto ES como MS conseguem se aproximar do desempenho final das demais estratégias, enquanto a LC permanece inferior ao longo de toda a análise.

4.1.3 Resultados em Parasitos

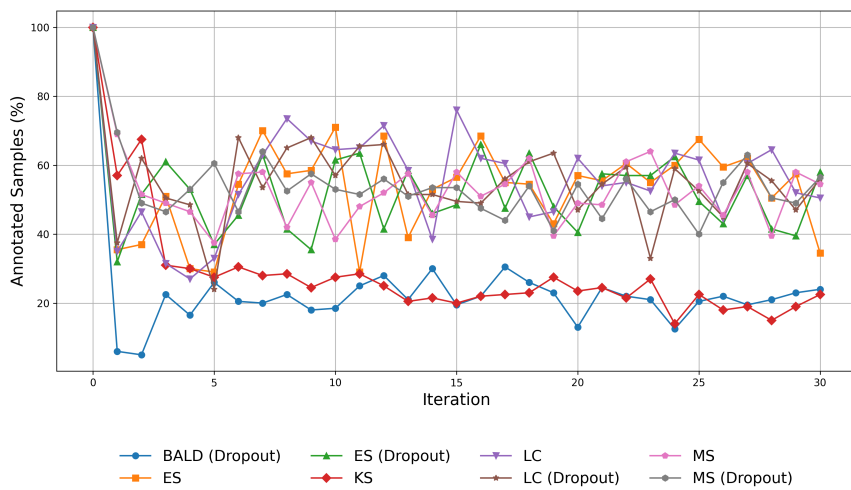
A Figura 8(a) apresenta a evolução da acurácia no conjunto de dados Parasitos e evidencia um comportamento mais desafiador do que os observados no MNIST e no FashionMNIST, refletindo a maior complexidade visual, o desbalanceamento entre classes e a maior variabilidade intra-classe do conjunto. O modelo supervisionado completo alcança cerca de 94,94% de acurácia, utilizando 100% (41332 amostras) do conjunto de dados. As estratégias de aprendizado ativo convergem para valores próximos, com uma redução de dados anotados bastante significativa. Por exemplo, a estratégia MS (Dropout) atinge



(a)



(b)



(c)

Figura 7 – Resultados obtidos por cada estratégia de aprendizado ativo utilizando o conjunto FashionMNIST. (a) acurácias. (b) quantidade de classes conhecidas e (c) porcentagem de amostras corrigidas

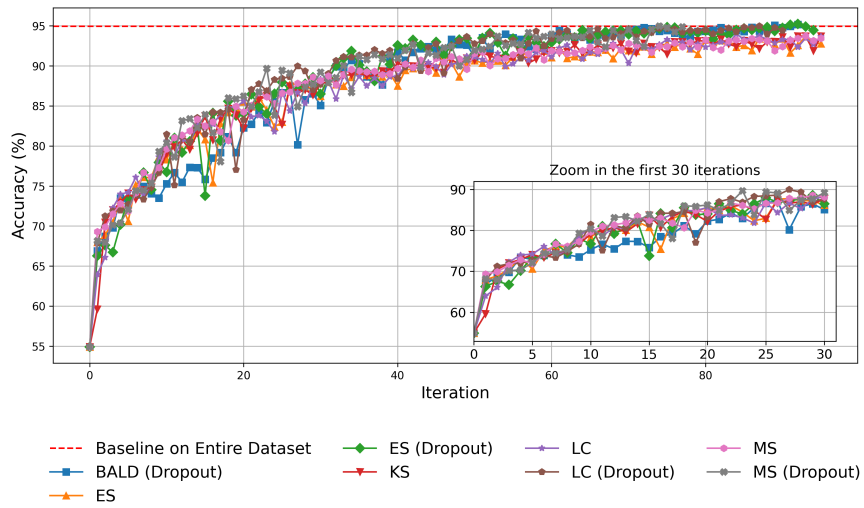
cerca de 95% de acurácia com apenas 36.28% (15000 amostras) do conjunto (Iteração 74 na Figura 8(a)). Analisando os desempenhos entre as estratégias de aprendizado ativo, observa-se que a *KMeans Sampling* (KS) apresenta leve instabilidade ao longo das iterações, com quedas pontuais em sua curva de acurácia, enquanto *Least Confidence* (LC), *Entropy Sampling* (ES) e *Margin Sampling* (MS) mantêm trajetórias mais regulares e estáveis.

Diferentemente dos conjuntos de dados MNIST e FashionMNIST, no conjunto de dados Parasitos (Figura 8(b)), nenhuma estratégia de aprendizado ativo apresenta o conhecimento de todas as (18) classes. A estratégia de aprendizado MS seleciona amostras de todas as classes mais rapidamente (na terceira iteração) em relação às demais estratégias. Analisando a quantidade de amostras corrigidas (Figura 8(c)), pode-se observar que as estratégias baseadas em Dropout tendem a corrigir um volume ligeiramente maior de amostras.

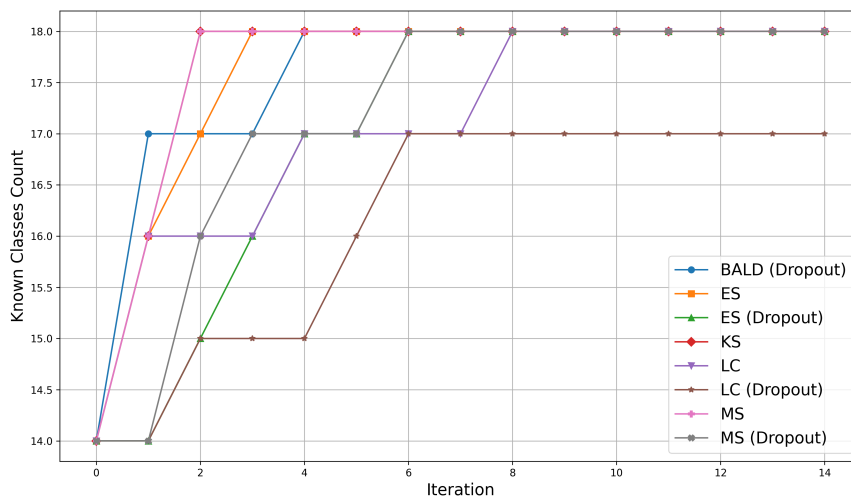
Análise Geral das Estratégias Individuais

A análise integrada dos três conjuntos de dados evidencia que o comportamento das estratégias individuais de Aprendizado Ativo varia conforme a complexidade visual, o número de classes e o grau de separabilidade entre categorias em cada domínio. Em bases simples e bem estruturadas, como o MNIST, todas as estratégias demonstram rápida evolução nas primeiras iterações, aproximando-se do desempenho do modelo totalmente supervisionado com apenas uma fração do conjunto rotulado. Nesse caso, abordagens baseadas em diversidade e margem (KS e MS) aceleram o aprendizado inicial, enquanto critérios de incerteza (LC e ES) alcançam os melhores resultados finais, evidenciando uma relação complementar entre representatividade e refinamento da fronteira de decisão. No FashionMNIST, embora KS e MS mantenham vantagem inicial, os métodos baseados em incerteza apresentam progressão mais lenta e, no caso da LC, desempenho inferior durante quase todo o processo. Já no Parasitos, onde há forte desbalanceamento e grande quantidade de classes, as diferenças entre estratégias tornam-se ainda menos expressivas em termos de acurácia, e todas apresentam evolução mais gradual, refletindo o maior desafio do domínio.

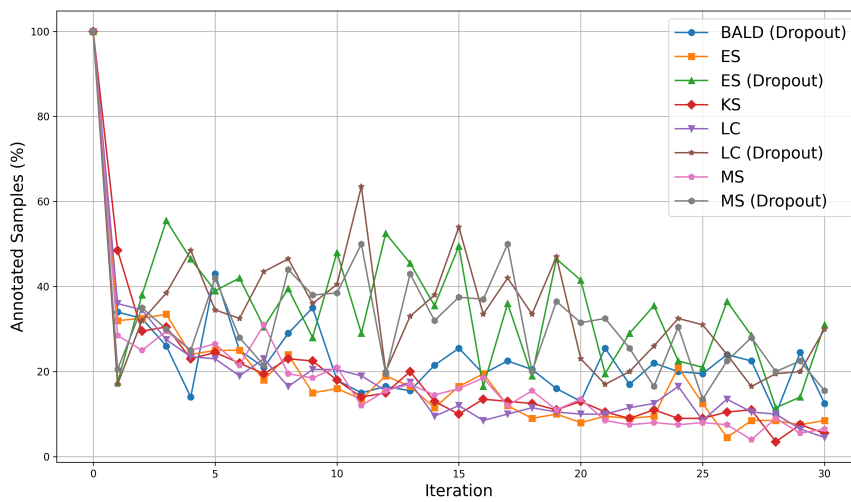
Outro aspecto relevante refere-se à quantidade de correções realizadas nas amostras selecionadas. O KS tende a selecionar exemplos mais representativos e pouco ambíguos, resultando em menor necessidade de correção e maior estabilidade preditiva. Em contraste, MS e, em menor grau, LC e ES, trazem para o conjunto rotulado instâncias mais próximas da fronteira de decisão, o que aumenta o número de ajustes ao longo das iterações. Tal padrão reforça a ideia de que diversidade e incerteza capturam diferentes regiões do espaço de características e contribuem de forma distinta para o processo de aprendizado. Em síntese, os resultados demonstram que o Aprendizado Ativo é altamente eficaz em reduzir o esforço de anotação, obtendo-se desempenho competitivo com apenas uma fração das



(a)



(b)



(c)

Figura 8 – Resultados obtidos por cada estratégia de aprendizado ativo utilizando o conjunto Parasitos. (a) acurácias. (b) quantidade de classes conhecidas e (c) porcentagem de amostras corrigidas

amostras rotuladas. Além disso, os achados confirmam que os critérios de seleção atuam de maneira complementar: enquanto a diversidade favorece a cobertura inicial do espaço e acelera o aprendizado no começo, técnicas baseadas em incerteza se tornam mais efetivas quando o modelo já consolidou sua estrutura decisória. Essa complementaridade motiva, de forma direta, a investigação realizada na Seção 4.2, na qual estratégias cooperativas baseadas em ensembles são exploradas a fim de superar as limitações observadas nas abordagens individuais e obter um processo de seleção mais robusto em diferentes cenários.

4.2 Cenário 2 – Resultados das Combinações das Estratégias de Aprendizado Ativo

A segunda etapa dos experimentos avalia o comportamento de combinações entre as estratégias de Aprendizado Ativo, buscando compreender se a integração de critérios distintos, como diversidade e incerteza, produz efeitos complementares capazes de superar as limitações observadas nas abordagens individuais. Para isso, cada estratégia (EN, KS, LC, MS) foi combinada par a par, e cada par foi analisado em quatro cenários distintos. No primeiro cenário, uma estratégia é aplicada ao conjunto completo para selecionar um subconjunto inicial de amostras; em seguida, a segunda estratégia é aplicada apenas sobre esse subconjunto, resultando em uma seleção ainda mais refinada. O segundo cenário é análogo, porém invertendo-se a ordem de aplicação das estratégias. Já o terceiro e o quarto cenários referem-se, respectivamente, à intersecção e à união das amostras selecionadas por ambas as estratégias quando aplicadas de forma independente. Cada combinação foi analisada considerando a acurácia, a porcentagem de amostras corrigidas, a quantidade de classes conhecidas (Apêndices A-C), bem como os tempos de seleção, treinamento e teste, permitindo uma análise abrangente do impacto de cada forma de integração sobre o processo de seleção ativa.

4.2.1 Resultados em MNIST

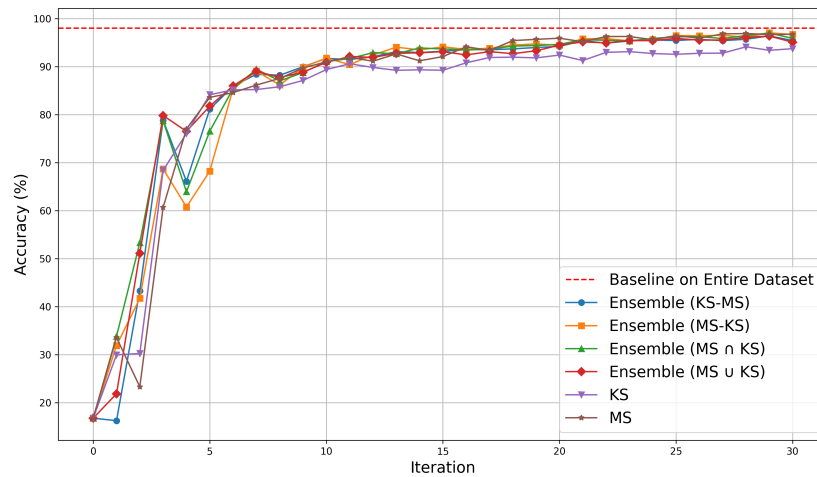
O aprendizado ativo baseado em ensemble em MNIST é altamente eficaz, atingindo resultados próximos ao treinamento supervisionado completo com poucas iterações. Pode-se observar pela Figura 9 que, em todos os cenários avaliados, as combinações entre estratégias resultam em melhorias em relação ao desempenho obtido pelo uso individual. Após cerca de 10 iterações, a diferença entre ensembles e modelos individuais reduz, indicando uma convergência similar no longo prazo. No entanto, os ensembles continuam ligeiramente superiores e mais estáveis. Nos cenários 1 e 2, em que uma estratégia é aplicada sequencialmente após a outra, os resultados indicam, de forma geral, que iniciar o processo de seleção com uma abordagem baseada em diversidade e, posteriormente, aplicar um critério baseado em incerteza tende a produzir desempenhos superiores ao

cenário inverso. Esse padrão sugere que a diversidade contribui para uma cobertura inicial mais ampla do espaço de dados, enquanto a incerteza atua de maneira mais eficaz no refinamento das regiões decisórias à medida que o modelo evolui.

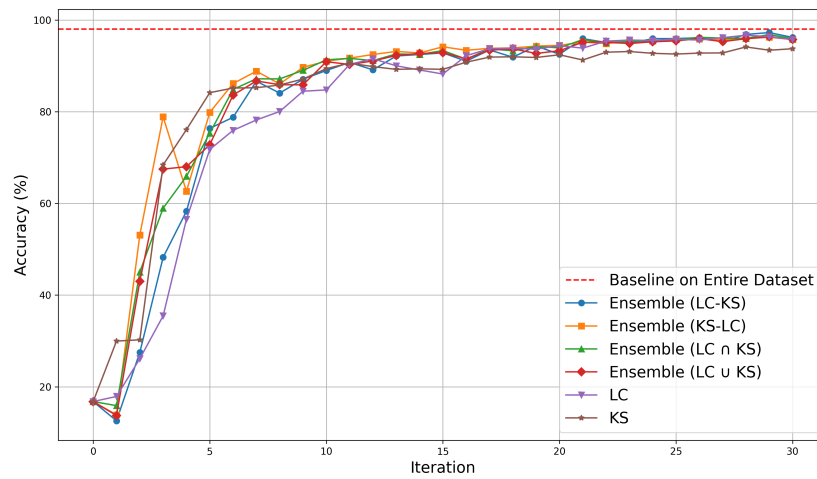
Entre todas as combinações avaliadas, de maneira geral, além do cenário explorando diversidade seguida de incerteza, destaca-se também o cenário correspondente à união das amostras selecionadas por cada par de estratégias, que apresentou alguns dos melhores resultados. Esse desempenho sugere que cada método captura aspectos complementares da estrutura do conjunto de dados, ampliando o conjunto de instâncias informativas disponíveis ao modelo. De modo geral, esses achados reforçam o potencial dos comitês de aprendizado ativo em combinar critérios distintos para maximizar a eficiência do processo de anotação e acelerar o ganho de desempenho.

Em todos os gráficos de radar (Figura 10), observa-se um padrão consistente: as combinações em comitê tendem a manter acurácia de teste igual ou ligeiramente superior às melhores estratégias individuais do par. A porcentagem de erros acompanha esse comportamento (quando a acurácia cresce, a fatia de erros tende a diminuir). Em praticamente todos os casos, todas as abordagens conhecem todas as classes ao final do processo, indicando que, em MNIST, a dificuldade não está em descobrir novas classes, mas em refinar a fronteira de decisão. Nas combinações KS-ES e LC-ES, a variante intersecção aparece com acurácia entre as mais altas do grupo, sugerindo que exigir consenso entre os critérios (diversidade e incerteza) ajuda a filtrar amostras mais informativas. Nas combinações com MS, a própria MS isolada já apresenta desempenho muito forte em MNIST; os ensembles mantêm acurácia competitiva, mas nem sempre superam de forma clara a estratégia individual, o que está alinhado ao fato de MNIST ser um problema relativamente simples.

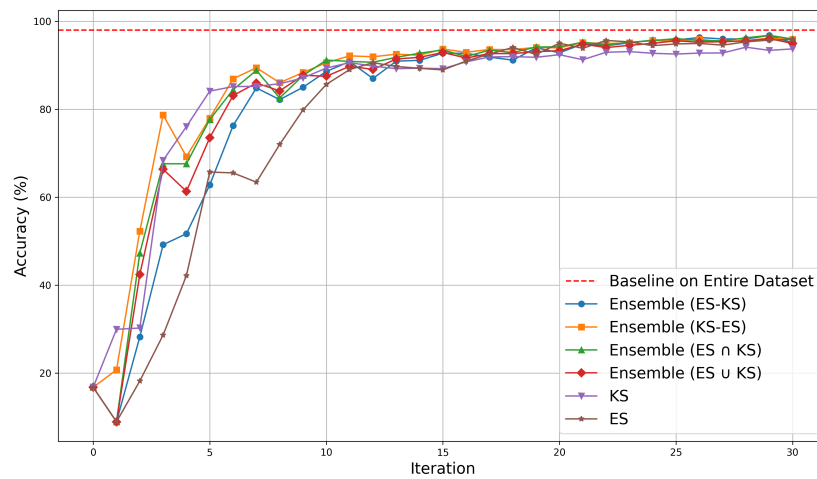
Os radares também deixam claro o impacto das combinações no custo computacional: as variantes que envolvem união ou intersecção tendem a apresentar maior tempo de seleção, pois exigem a execução de dois critérios de AL de forma independente e, em seguida, a combinação dos subconjuntos. Os cenários sequenciais (por exemplo, KS→ES ou ES→KS) costumam ter tempo de seleção intermediário, já que a segunda estratégia atua sobre um subconjunto reduzido. O tempo de treinamento e teste é, em geral, similar entre estratégias que usam o mesmo número de iterações e de amostras rotuladas, com pequenas variações decorrentes das diferenças no conjunto selecionado. Nos gráficos, isso aparece como vértices relativamente próximos para a maior parte dos métodos, com uma ou outra combinação sendo ligeiramente mais onerosa. Em resumo, os ensembles oferecem ganhos preditivos modestos em MNIST, mas à custa de um aumento de tempo de seleção, especialmente nas variantes de união/intersecção.



(a)



(b)



(c)

Figura 9 – Resultados de acurácias obtidos pelo uso individual e pela combinação da estratégia de aprendizado ativo baseada em diversidade (KS) com as demais estratégias baseadas em incerteza: (a) MS, (b) LC e (c) EN, respectivamente, utilizando o conjunto de dados MNIST.

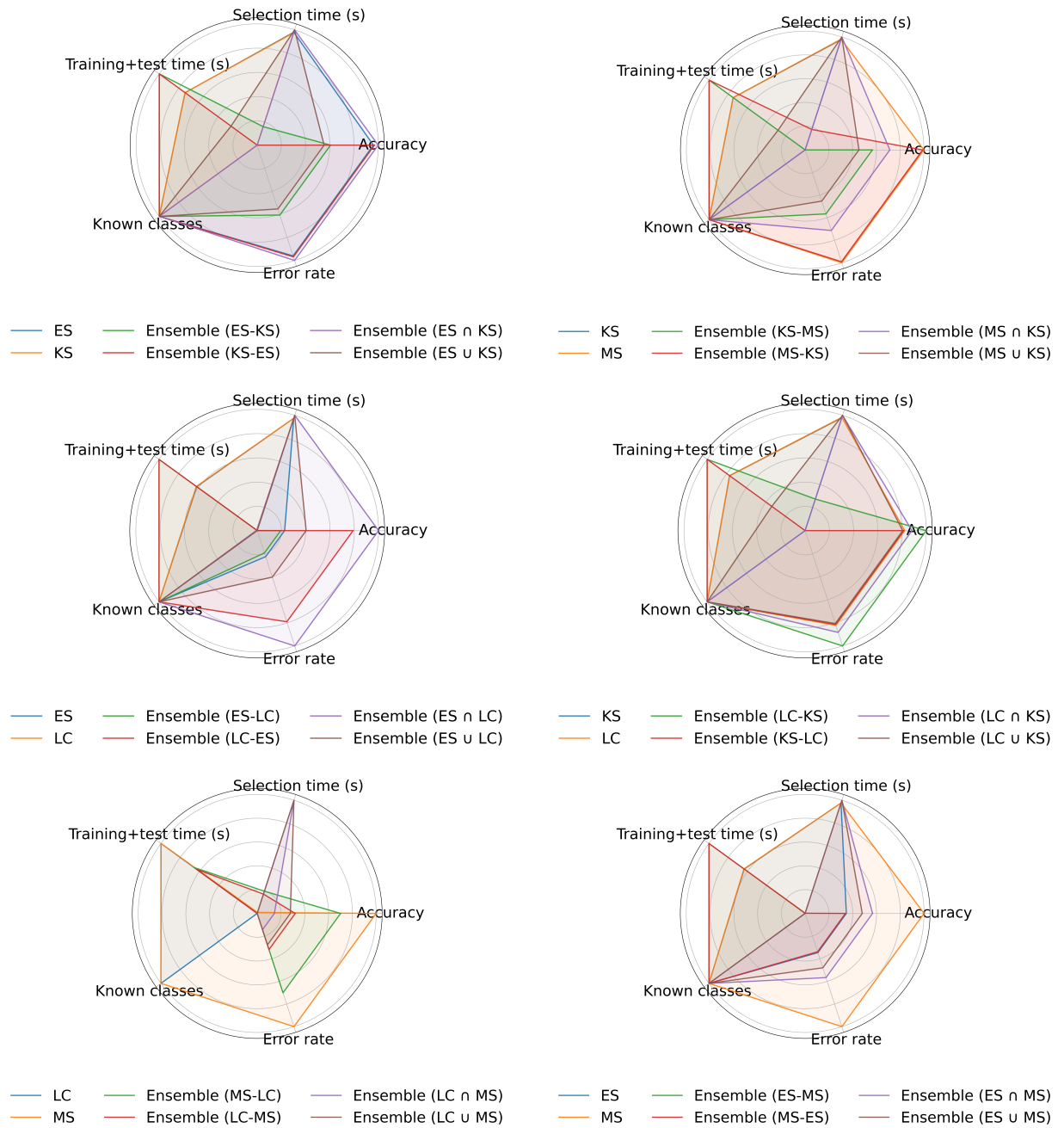


Figura 10 – Gráficos de radar sintetizando o desempenho das estratégias individuais e de suas combinações em termos de acurácia no teste, porcentagem de erros, número de classes conhecidas, tempo de seleção e tempo de treino+teste, utilizando o conjunto de dados MNIST. Cada subfigura compara um par de estratégias de aprendizado ativo e suas variantes em comitê: (a) KS-ES, (b) KS-MS, (c) LC-ES, (d) LC-KS, (e) LC-MS e (f) MS-ES. Todas as métricas foram normalizadas no intervalo $[0, 1]$ para fins de comparação relativa.

4.2.2 Resultados em FashionMNIST

A avaliação dos métodos de seleção incremental de amostras no conjunto de dados FashionMNIST revela tendências consistentes entre as diferentes estratégias testadas, bem como vantagens claras no uso de abordagens baseadas em ensembles (Figura 11). Em

todos os experimentos, observa-se um comportamento típico de cenários de aprendizado ativo: as acurácias iniciais partem de valores baixos (10–20%), crescem rapidamente nas primeiras iterações devido à incorporação de amostras informativas e estabilizam-se após aproximadamente dez iterações, atingindo valores entre 65% e 75%. Entre as técnicas individuais avaliadas, o método KS destaca-se como o mais consistente e robusto. Sua estratégia de maximização da diversidade entre as amostras selecionadas favorece boa cobertura do espaço de características, resultando em curvas mais estáveis e acurácias superiores às obtidas por métodos puramente baseados em incerteza. Em contraste, a técnica Least Confidence (LC) apresenta o pior desempenho individual. Por considerar apenas a confiança do modelo sem levar em conta diversidade ou complexidade do espaço de entrada, LC tende a escolher amostras redundantes, o que limita a capacidade de generalização e explica seu crescimento mais lento ao longo das iterações. As técnicas Entropy Sampling (ES) e Margin Sampling (MS) apresentam desempenhos intermediários, geralmente superiores ao LC, mas ainda inferiores ao KS. ES explora diretamente a incerteza do modelo, enquanto MS enfatiza amostras próximas à fronteira de decisão. Ambas as abordagens contribuem para selecionar exemplos relevantes, mas podem carecer de diversidade suficiente quando aplicadas isoladamente.

Os resultados tornam-se mais expressivos quando se analisam os métodos baseados em ensembles, que combinam propriedades de diferentes estratégias de seleção. De maneira geral, os ensembles superam as técnicas individuais em praticamente todos os cenários avaliados, evidenciando que a combinação de critérios produz conjuntos de amostras mais equilibrados e representativos. Entre todas as variantes, destaca-se o Ensemble Union, que, ao considerar a união das amostras selecionadas por mais de uma técnica, maximiza simultaneamente diversidade e incerteza. Isso resulta em desempenho superior e maior estabilidade ao longo das iterações. Ao comparar os três cenários principais, combinações KS-MS, KS-LC, KS-ES (Figuras 11(a)-11(c), respectivamente), observa-se que os ensembles baseados em KS com métodos de incerteza (ES ou MS) apresentam desempenhos mais elevados. Em particular, o cenário KS-MS demonstra os melhores resultados gerais, atingindo valores próximos de 74–75% e mantendo estabilidade ao longo das iterações. Esse comportamento indica que a combinação entre diversidade (KS) e incerteza estruturada (MS) produz amostras especialmente informativas para o processo de aprendizado incremental.

De modo geral, os gráficos de radar para o FashionMNIST (Figura 12) evidenciam o impacto positivo do uso de ensembles de aprendizado ativo. As combinações entre diversidade e incerteza resultam em melhorias consistentes sobre as abordagens individuais, reduzindo a redundância das amostras escolhidas e acelerando tanto o ganho de acurácia quanto a descoberta das classes. Destacam-se especialmente as combinações envolvendo o KS, que se mostraram mais equilibradas entre desempenho preditivo, custo computacional e cobertura do espaço amostral. Entre as estratégias de fusão avaliadas, a união tende a

proporcionar os maiores ganhos de desempenho, enquanto a intersecção apresenta maior estabilidade e economia de seleção. Assim, conclui-se que ensembles são particularmente vantajosos em cenários intermediários de complexidade, como o FashionMNIST, nos quais a complementaridade entre diversidade e incerteza torna o processo de seleção ativa mais eficiente e robusto.

4.2.3 Resultados em Parasitos

No conjunto de dados Parasitos, todos os métodos de aprendizado ativo baseados em ensemble melhoram progressivamente a acurácia ao longo das iterações (Figura 13). Observa-se que os ensembles superam, de forma sistemática, os modelos individuais, evidenciando o benefício da colaboração entre diferentes classificadores na seleção de amostras informativas. Entre as combinações, o KS destaca-se, apresentando melhores resultados finais (todas as possíveis combinações entre as estratégias de aprendizado ativo encontram-se disponíveis no Apêndice C). Entre as estratégias de fusão, intersecção e união, a intersecção demonstra maior estabilidade e alta precisão, enquanto que a união apresenta desempenho competitivo, porém com maior variabilidade ao longo das iterações. Sendo assim, o emprego de aprendizado ativo com ensemble mostra-se eficiente ao acelerar a obtenção de bons níveis de acurácia e reduzir a dependência de grandes volumes de rótulos. Em particular, a combinação do KS com a estratégia ensemble baseada em intersecção alcançou o melhor desempenho no conjunto de dados de parasitos.

No conjunto de dados Parasitos, os gráficos de radar (Figura 14) evidenciam que o uso de ensembles torna o aprendizado ativo significativamente mais eficiente, promovendo simultaneamente maior acurácia, menor taxa de erros e descoberta mais rápida das classes. O ganho de desempenho está diretamente associado à complementaridade entre diversidade e incerteza: métodos individuais apresentam avanços limitados e maior variabilidade, enquanto os ensembles, especialmente as combinações envolvendo o KS com critérios de incerteza como ES, mostram melhor equilíbrio entre cobertura do espaço amostral, estabilidade de seleção e custo computacional. Entre as estratégias de fusão, a união tende a alcançar a maior acurácia quando há maior disponibilidade de rótulos, ao passo que a intersecção é mais eficiente e consistente em cenários mais restritivos, reduzindo redundâncias e tempo de seleção. Assim, os resultados demonstram que a integração de múltiplos critérios de aquisição melhora todas as métricas simultaneamente, sendo particularmente vantajosa em contextos complexos como o do Parasitos.

Análise Geral das Combinações

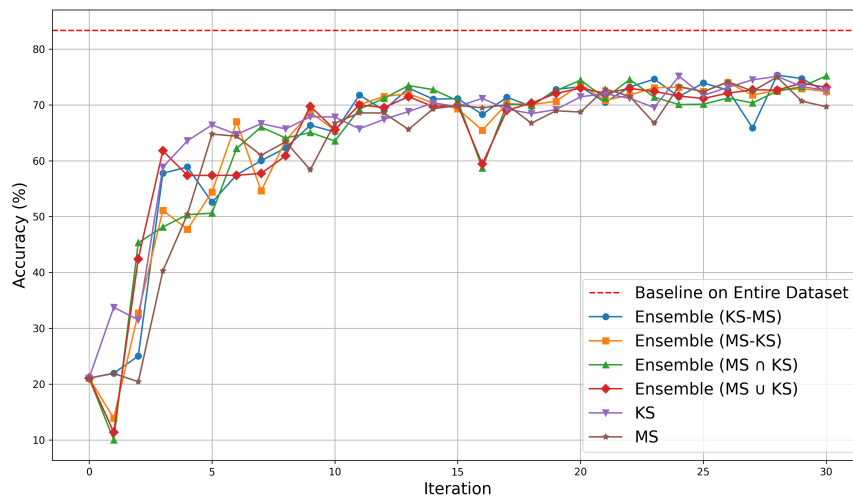
De forma integrada, os resultados obtidos nos três conjuntos de dados indicam que o impacto das combinações de estratégias de Aprendizado Ativo depende fortemente da complexidade do domínio e da forma como os critérios são articulados em cada

cenário de ensemble. Em bases simples e bem separadas, como o MNIST, as estratégias individuais já são capazes de atingir rapidamente acurácias próximas ao modelo totalmente supervisionado, o que limita o espaço de ganho das combinações. Nesses casos, os ensembles tendem a reproduzir o comportamento da melhor estratégia do par, produzindo melhorias modestas, sobretudo em termos de estabilidade das curvas, mas com ganhos marginais de acurácia e pouca diferença entre os cenários de intersecção e união.

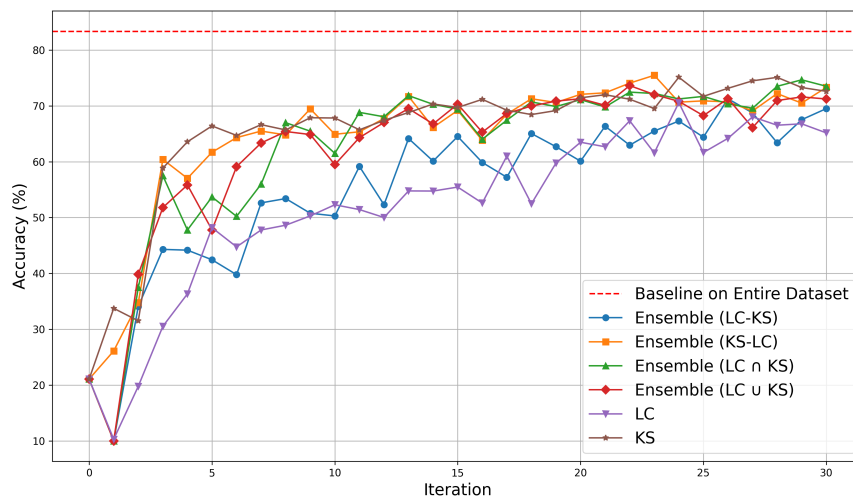
À medida que a complexidade do problema aumenta, como em FashionMNIST e, de forma ainda mais evidente, em Parasitos, o papel dos ensembles torna-se mais expressivo. Nestes domínios, a maior variabilidade intra-classe, a sobreposição entre categorias e, no caso de Parasitos, o desbalanceamento entre classes, ampliam o benefício da integração de critérios distintos. Observa-se que combinações que articulam diversidade e incerteza tendem a produzir conjuntos de amostras rotuladas mais equilibrados, resultando em curvas de aprendizado mais estáveis e, em vários casos, em acurácias superiores às obtidas pelas estratégias isoladas. Esse efeito é particularmente claro em combinações envolvendo a KS com métodos baseados em incerteza (ES, LC e MS), nas quais tanto os cenários sequenciais quanto os de união/intersecção frequentemente superam as abordagens individuais.

Os cenários em que uma estratégia baseada em diversidade é aplicada inicialmente, seguida por um critério de incerteza, mostraram-se especialmente promissores. Esse arranjo explora, nas iterações iniciais, a capacidade da diversidade de promover uma cobertura mais ampla do espaço de características, permitindo que o conjunto rotulado represente melhor as diferentes regiões do domínio. Em seguida, a incerteza atua de forma mais eficaz no refinamento local da fronteira de decisão, selecionando instâncias mais ambíguas ou difíceis à medida que o modelo amadurece. De modo complementar, os cenários baseados em união tendem a explorar de forma mais agressiva a complementaridade entre estratégias, reunindo amostras selecionadas por critérios distintos e ampliando o conjunto de instâncias potencialmente informativas. Já a intersecção, embora mais conservadora, mostrou-se particularmente útil para reduzir oscilações e eliminar amostras sobre as quais não há consenso entre os critérios, contribuindo para maior estabilidade preditiva, como observado no conjunto Parasitos.

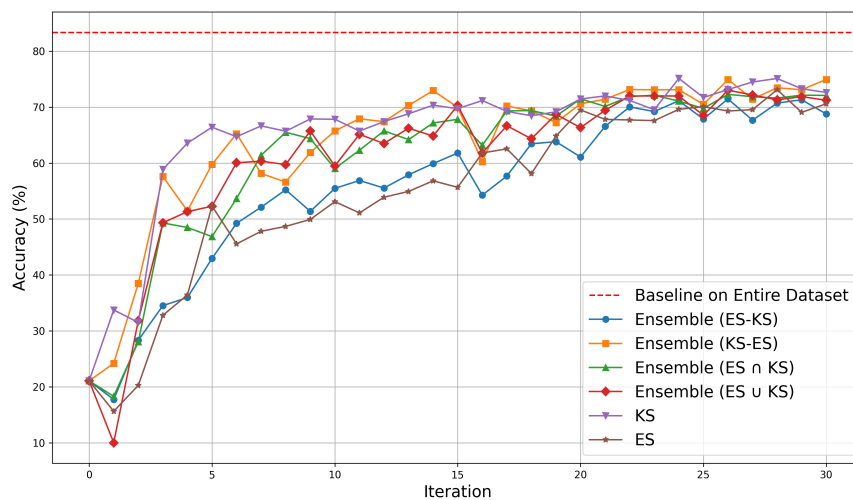
Em síntese, os resultados indicam que, embora os ganhos dos ensembles sejam limitados em domínios de baixa complexidade, sua contribuição torna-se significativa em cenários mais desafiadores, nos quais a combinação de diversidade e incerteza permite explorar o espaço de entrada de forma mais equilibrada. Esses achados reforçam a hipótese central deste trabalho de que comitês de estratégias de Aprendizado Ativo podem mitigar limitações das abordagens individuais, acelerar a convergência e melhorar a estabilidade do processo de seleção de amostras, especialmente em bases com alta variabilidade visual, grande número de classes ou desbalanceamento acentuado.



(a)



(b)



(c)

Figura 11 – Resultados de acurárias obtidos pelo uso individual e pela combinação da estratégia de aprendizado ativo baseada em diversidade (KS) com as demais estratégias baseadas em incerteza: (a) MS, (b) LC, (c) EN, respectivamente, utilizando o conjunto de dados FashionMNIST.

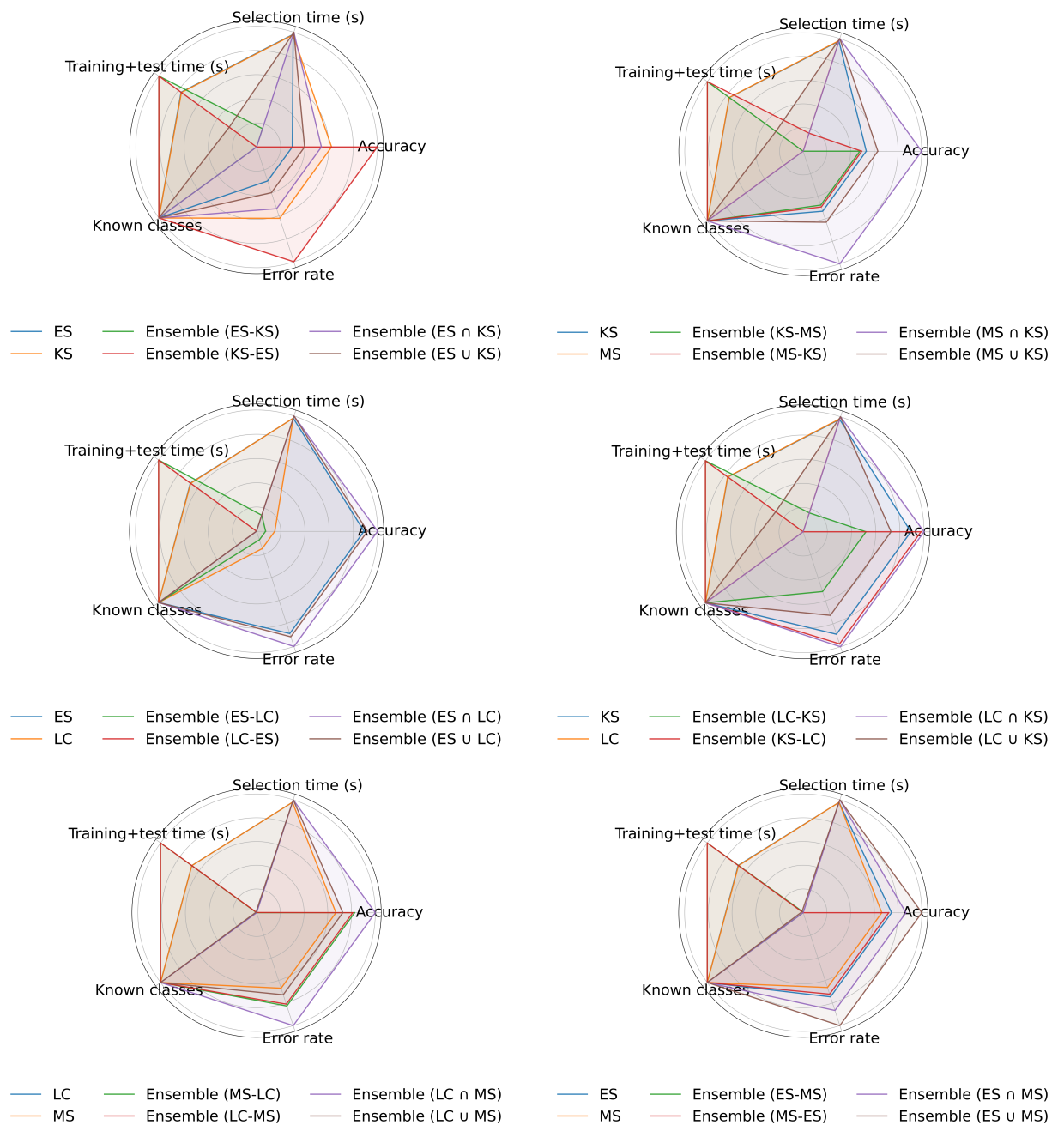
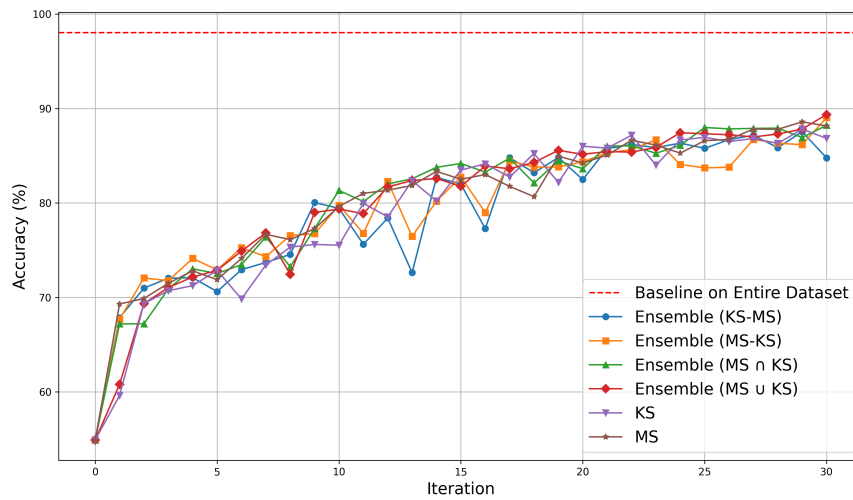
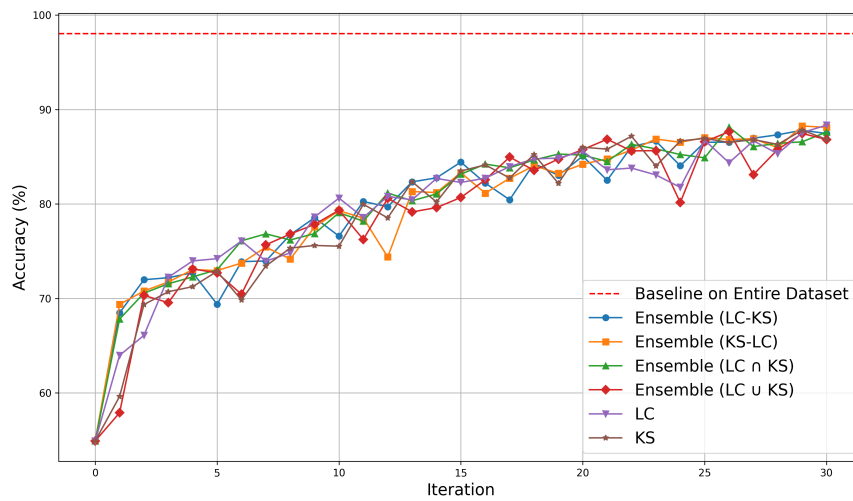


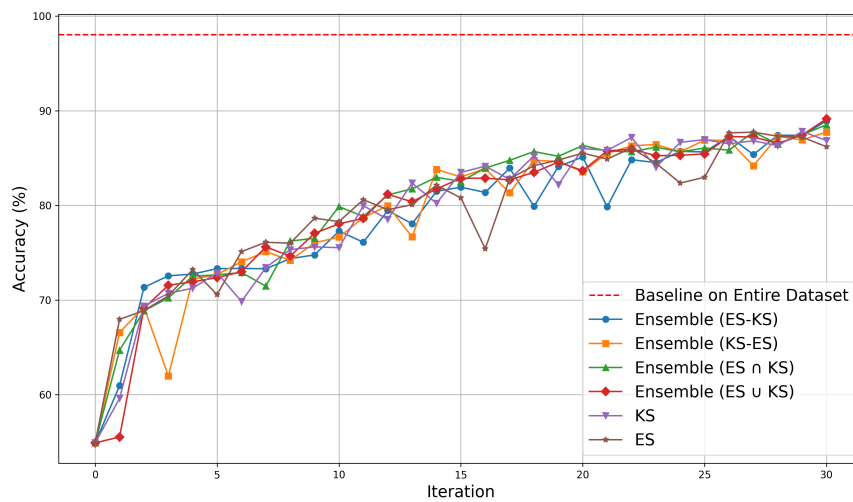
Figura 12 – Gráficos de radar sintetizando o desempenho das estratégias individuais e de suas combinações em termos de acurácia no teste, porcentagem de erros, número de classes conhecidas, tempo de seleção e tempo de treino+teste, utilizando o conjunto de dados FashionMNIST. Cada subfigura compara um par de estratégias de aprendizado ativo e suas variantes em comitê: (a) KS-ES, (b) KS-MS, (c) LC-ES, (d) LC-KS, (e) LC-MS e (f) MS-ES. Todas as métricas foram normalizadas no intervalo [0, 1] para fins de comparação relativa.



(a)



(b)



(c)

Figura 13 – Resultados de acurácias obtidos pelo uso individual e pela combinação da estratégia de aprendizado ativo baseada em diversidade (KS) com as demais estratégias baseadas em incerteza: (a) MS, (b) LC, (c) EN, respectivamente, utilizando o conjunto de dados Parasitos.

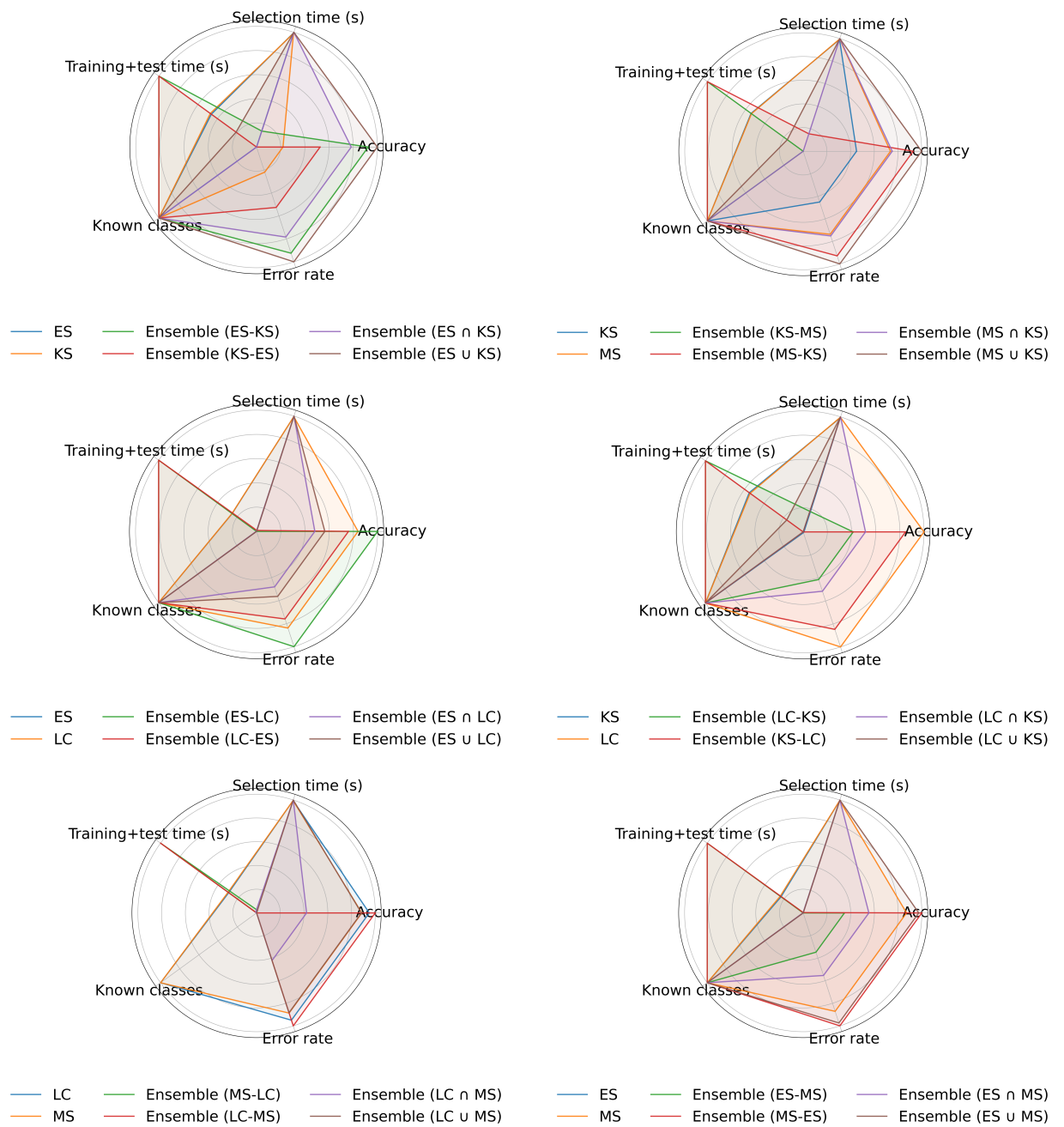


Figura 14 – Gráficos de radar sintetizando o desempenho das estratégias individuais e de suas combinações em termos de acurácia no teste, porcentagem de erros, número de classes conhecidas, tempo de seleção e tempo de treino+teste, utilizando o conjunto de dados Parasitos. Cada subfigura compara um par de estratégias de aprendizado ativo e suas variantes em comitê: (a) KS-ES, (b) KS-MS, (c) LC-ES, (d) LC-KS, (e) LC-MS e (f) MS-ES. Todas as métricas foram normalizadas no intervalo $[0, 1]$ para fins de comparação relativa.

5 Conclusão

Este capítulo consolida os principais achados e implicações deste trabalho, retomando os objetivos apresentados na Introdução e discutindo a contribuição efetiva das combinações de estratégias de aprendizado ativo profundo para a seleção eficiente de amostras. Ao longo da pesquisa, buscou-se compreender em que medida a integração entre critérios de incerteza e diversidade poderia aprimorar a performance de modelos neurais em cenários com restrição de rótulos, reforçando tanto o potencial quanto os limites do uso de comitês em tarefas de classificação visual.

O objetivo central consistiu em propor, implementar e validar uma abordagem de aprendizado ativo baseada em comitês, capaz de ponderar diferentes perspectivas sobre a informatividade das instâncias, ora privilegiando a incerteza preditiva, ora explorando a diversidade estrutural dos dados. A partir desse propósito, desenvolveu-se uma infraestrutura experimental reproduzível, contemplando a implementação das estratégias *Least Confidence*, *Margin Sampling*, *Entropy Sampling* e *KMeans Sampling*, bem como suas combinações por sequência, união e intersecção. O protocolo experimental aplicado aos conjuntos MNIST, FashionMNIST e Parasitos permitiu avaliar, de forma sistemática, o comportamento das técnicas em cenários de crescente complexidade.

Os resultados obtidos mostram que o impacto dos comitês depende fortemente das características do domínio. Em conjuntos de dados simples, como o MNIST, os ganhos adicionais foram modestos: embora as combinações antecipem em algumas iterações a convergência para altos níveis de acurácia, a vantagem numérica permanece pequena. Ainda assim, o experimento confirma o papel fundamental do aprendizado ativo na redução do custo de anotação, uma vez que acurácias próximas às do modelo totalmente supervisionado foram atingidas com aproximadamente 17,5% dos rótulos disponíveis.

Um dos cenários mais desafiador, representado pelo FashionMNIST, reforçou a relevância das combinações entre incerteza e diversidade. Comitês como KS-MS e ES-KS alcançaram melhorias substanciais, superando as estratégias isoladas em até 5 p.p. e exibindo curvas mais estáveis ao longo das iterações. Esses resultados sugerem que, à medida que as classes tornam-se mais ambíguas e sobrepostas, abordagens híbridas passam a capturar nuances que critérios únicos não conseguem representar plenamente.

No conjunto Parasitos, de maior variabilidade intra-classe, os comitês apresentaram vantagens mais evidentes, principalmente na rapidez de descoberta de classes e na estabilidade das curvas de acurácia. A combinação ES-LC, por exemplo, obteve ganhos relevantes em cenários intermediários, sugerindo que a complementaridade entre entropia e confiança mínima é mais expressiva quando há maior complexidade estrutural no espaço

de características.

Além da análise quantitativa, este trabalho ressalta a importância do especialista (ou oráculo) no processo iterativo de aprendizado ativo, visto que parte significativa do ganho ocorre nas iterações intermediárias, quando o modelo ainda enfrenta incertezas distribuídas de maneira desigual entre as classes. Também se observou que comitês baseados em intersecção tendem a favorecer a precisão, ao passo que uniões ampliam a diversidade das amostras selecionadas — um equilíbrio que depende diretamente das características do conjunto de dados.

Apesar dos resultados positivos, algumas limitações devem ser reconhecidas. Os experimentos foram conduzidos em conjuntos de dados de pequeno e médio porte, o que restringe a generalização imediata para cenários de larga escala. A arquitetura empregada, ainda que adequada ao escopo deste trabalho, é relativamente simples e pode não capturar toda a complexidade de bases contemporâneas. Além disso, não foram consideradas variações estocásticas em múltiplas execuções, que poderiam aprofundar a análise de estabilidade entre execuções.

Como perspectivas futuras, destaca-se a possibilidade de aplicar o mesmo protocolo em conjuntos maiores e mais complexos, como CIFAR-100 ou ImageNet, além de investigar o uso de arquiteturas heterogêneas em comitês, incorporando redes do tipo ResNet e Vision Transformers. Outras extensões incluem o uso de técnicas semi-supervisionadas e auto-supervisionadas integradas ao ciclo de aprendizado ativo, reduzindo ainda mais a dependência de anotações manuais, e a adoção de análises estatísticas formais para comparar estratégias e quantificar a robustez dos comitês frente à variabilidade dos dados.

Em síntese, o estudo alcança seus objetivos ao demonstrar que a combinação estruturada de estratégias de aprendizado ativo pode aprimorar, em diferentes níveis, tanto a eficiência de rotulação quanto o desempenho preditivo de modelos visuais. Ao disponibilizar todo o código em repositório público, o trabalho reforça seu caráter reproduzível e contribui para o avanço da área. Conclui-se que comitês de aprendizado ativo representam uma alternativa promissora para contextos reais em que rótulos são escassos ou caros, abrindo caminho para investigações mais amplas e aplicações práticas em domínios cada vez mais desafiadores.

Referências

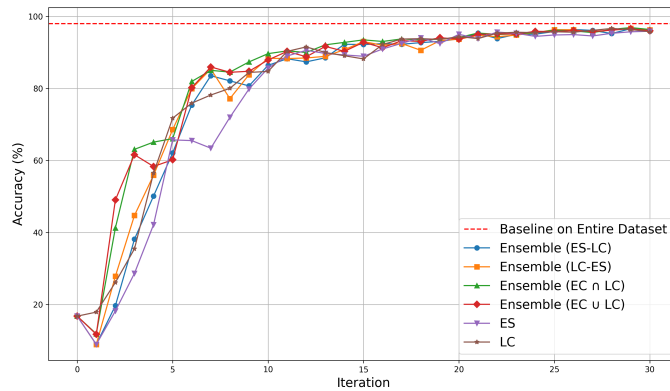
- CARRETE, J. et al. Deep ensembles vs committees for uncertainty estimation in neural-network force fields: Comparison and application to active learning. *The Journal of Chemical Physics*, AIP Publishing, v. 158, n. 20, maio 2023. ISSN 1089-7690. Disponível em: <<http://dx.doi.org/10.1063/5.0146905>>. 16, 18
- CHEN, Z. et al. Supervised anomaly detection via conditional generative adversarial network and ensemble active learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 45, n. 6, p. 7781–7798, 2023. 16, 18
- CHITTA, K. et al. Training data subset search with ensemble active learning. *IEEE Transactions on Intelligent Transportation Systems*, v. 23, n. 9, p. 14741–14752, 2022. 14, 16, 18
- DU, T.; ZHANG, S. Active learning with ensembles for doe. In: ZHANG, W. et al. (Ed.). *Current Trends in High Performance Computing and Its Applications*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005. p. 283–287. ISBN 978-3-540-27912-9. 14, 16, 17
- GAL, Y.; GHAHRAMANI, Z. Deep bayesian active learning with image data. In: *Proceedings of the 34th International Conference on Machine Learning*. [S.l.: s.n.], 2017. p. 1183–1192. 14, 17, 24
- HOULSBY, N. et al. Bayesian active learning for classification and preference learning. In: *Proceedings of the 28th International Conference on Machine Learning*. [S.l.: s.n.], 2011. p. 449–456. 14, 16, 17, 23
- HUANG, K.-H. Deepal: Deep active learning in python. *arXiv preprint arXiv:2111.15258*, 2021. 21, 24
- KRAWCZYK, B.; CANO, A. Adaptive ensemble active learning for drifting data stream mining. In: *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19*. International Joint Conferences on Artificial Intelligence Organization, 2019. p. 2763–2771. Disponível em: <<https://doi.org/10.24963/ijcai.2019/383>>. 14, 16, 18
- LECUN, Y.; CORTES, C.; BURGESS, C. Mnist handwritten digit database. *ATT Labs [Online]*. Available: <http://yann.lecun.com/exdb/mnist>, v. 2, 2010. 14, 20
- LEWIS, D. D.; GALE, W. A. A sequential algorithm for training text classifiers. In: *Proceedings of the 17th Annual International ACM SIGIR Conference*. [S.l.: s.n.], 1994. p. 3–12. 14, 16, 23
- NGUYEN, H.; SMEULDERS, A. Active learning using pre-clustering. In: *Proceedings of the 21st International Conference on Machine Learning*. [S.l.: s.n.], 2004. p. 79. 16, 17, 24
- POLYZOS, K. D.; LU, Q.; GIANNAKIS, G. B. Weighted ensembles for adaptive active learning. *IEEE Transactions on Signal Processing*, v. 72, p. 4178–4190, 2024. 14, 16, 18

- ROY, N.; MCCALLUM, A. Toward optimal active learning through monte carlo estimation of error reduction. In: *Proceedings of the 18th International Conference on Machine Learning*. [S.l.: s.n.], 2001. p. 441–448. 14, 16
- SETTLES, B. An analysis of active learning strategies for sequence labeling tasks. In: *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing*. [S.l.: s.n.], 2008. p. 1070–1079. 14, 16, 23
- WANG, H.; JIN, Y.; DOHERTY, J. Committee-based active learning for surrogate-assisted particle swarm optimization of expensive problems. *IEEE Transactions on Cybernetics*, v. 47, n. 9, p. 2664–2677, 2017. 16, 18
- XIAO, H.; RASUL, K.; VOLLGRAF, R. *Fashion-MNIST: a Novel Image Dataset for Benchmarking Machine Learning Algorithms*. 2017. Disponível em: <<https://arxiv.org/abs/1708.07747>>. 20

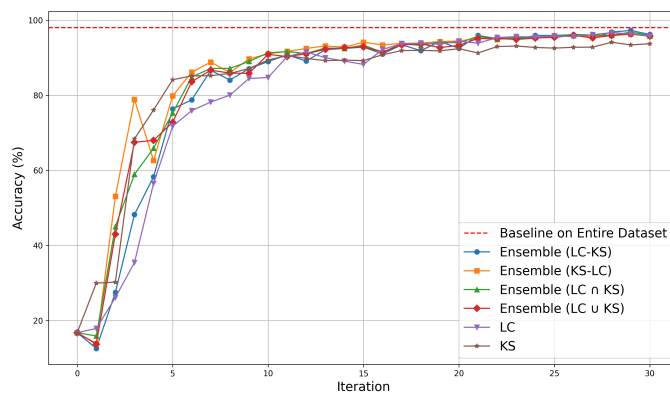
Apêndices

APÊNDICE A – MNIST

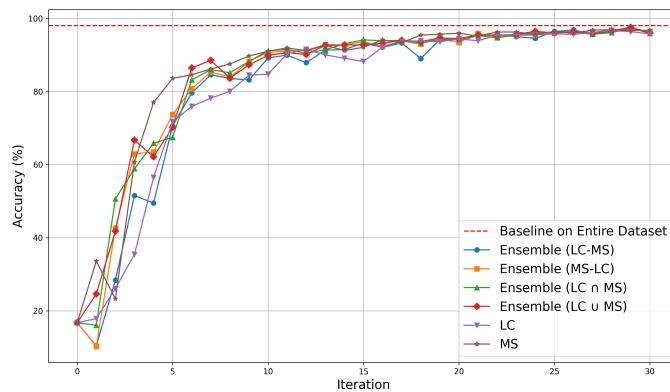
A.1 Combinações de Estratégias — Acurácia



(a)

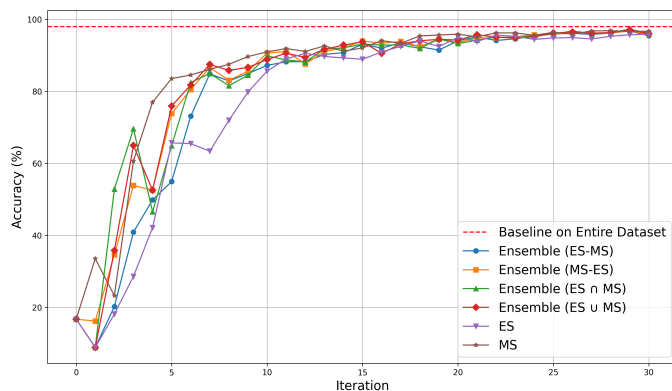


(b)

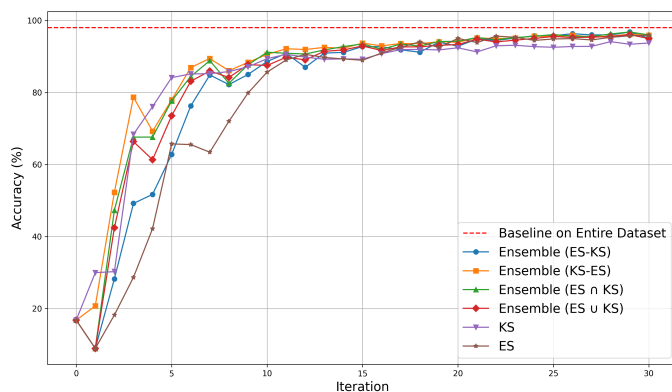


(c)

Figura 15 – Resultados de acurácias obtidos pelo uso individual e pela combinação da estratégia de aprendizado ativo baseada em incerteza (LC) com as demais estratégias baseadas em incerteza: (a) ES, (b) KS, (c) MS, respectivamente, utilizando o conjunto de dados MNIST



(a)



(b)

Figura 16 – Resultados de acurácia obtidos pela combinação da estratégia de aprendizado ativo baseada em entropia (ES) com as demais estratégias individuais: (a) MS e (b) KS, respectivamente, utilizando o conjunto de dados MNIST.

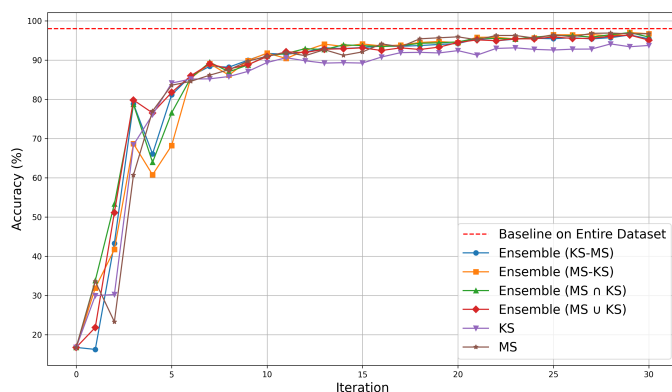
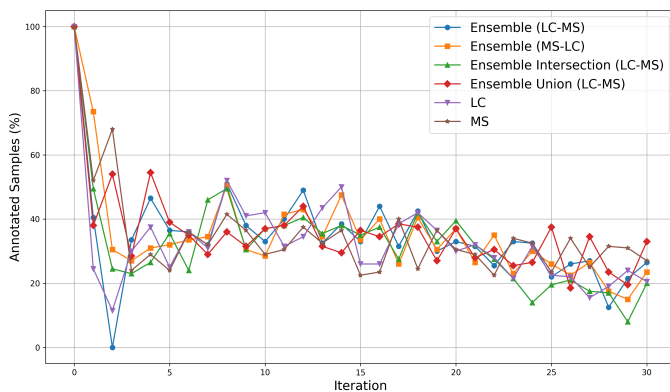
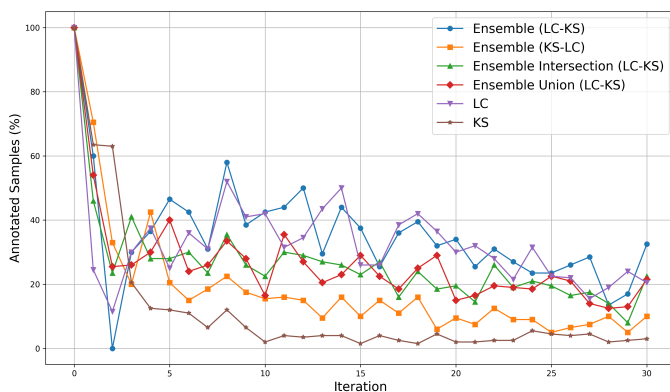


Figura 17 – Resultados de acurácia obtidos pela combinação da estratégia de aprendizado ativo baseada em margem (MS) com a estratégia baseada em diversidade (KS), utilizando o conjunto de dados MNIST.

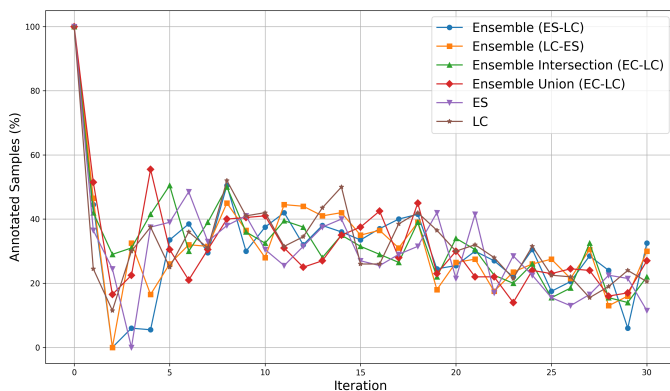
A.2 Combinações de Estratégias — Correções por Iterações



(a)

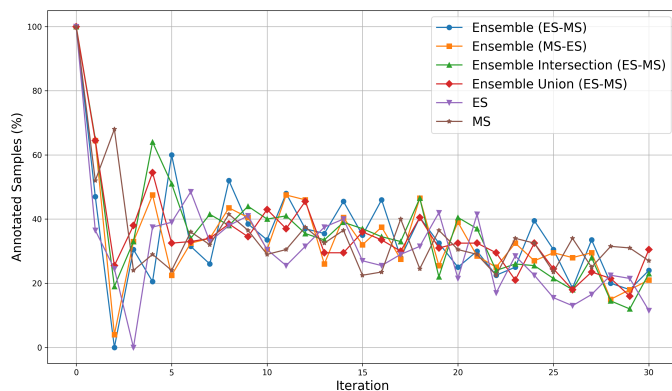


(b)

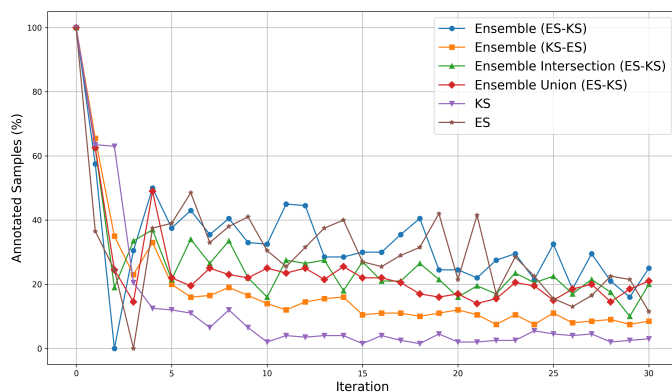


(c)

Figura 18 – Correções por iterações obtidas pela combinação da estratégia de aprendizado ativo baseada em incerteza (LC) com as estratégias individuais: (a) MS, (b) KS e (c) ES, respectivamente, utilizando o conjunto de dados MNIST.



(a)



(b)

Figura 19 – Correções por iterações obtidas pela combinação da estratégia de aprendizado ativo baseada em entropia (ES) com as estratégias individuais: (a) MS e (b) KS, respectivamente, utilizando o conjunto de dados MNIST.

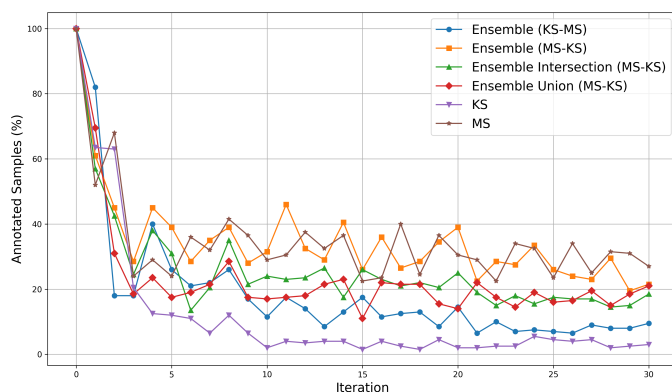
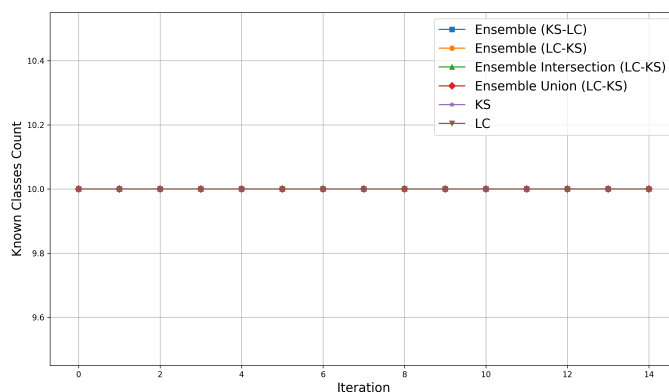
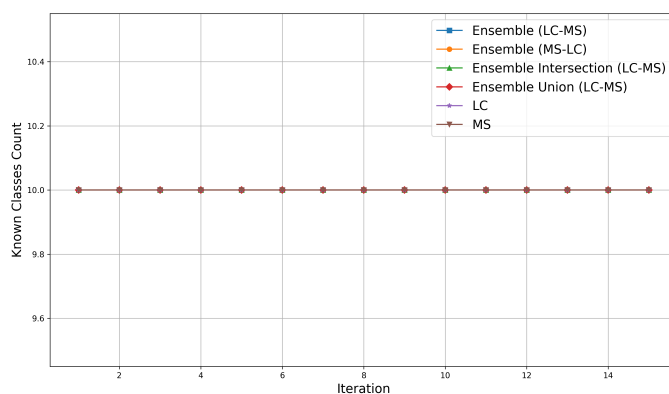


Figura 20 – Correção por iterações obtida pela combinação da estratégia de aprendizado ativo baseada em margem (MS) com a estratégia KS, utilizando o conjunto de dados MNIST.

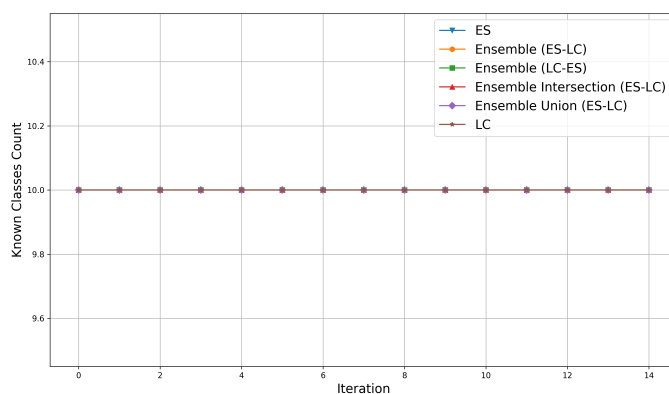
A.3 Combinações de Estratégias — Classes Conhecidas



(a)

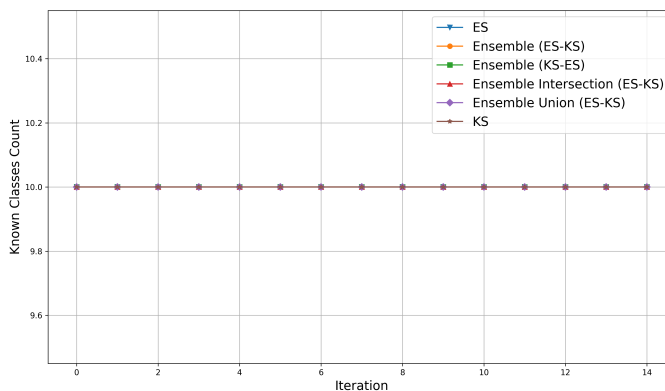


(b)

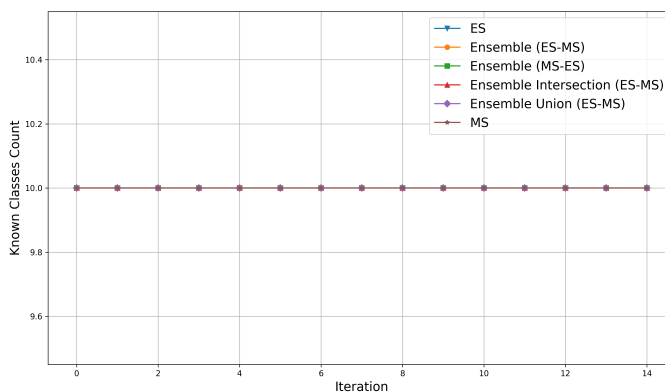


(c)

Figura 21 – Classes conhecidas ao longo das iterações utilizando a estratégia de aprendizado ativo baseada em incerteza (LC) combinada com: (a) KS, (b) MS e (c) ES no conjunto MNIST.



(a)



(b)

Figura 22 – Classes conhecidas ao longo das iterações utilizando a estratégia baseada em entropia (ES) combinada com: (a) KS e (b) MS no conjunto MNIST.

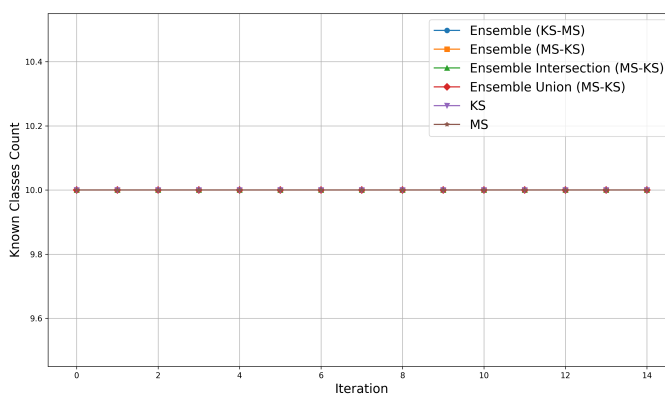


Figura 23 – Classes conhecidas ao longo das iterações utilizando a estratégia baseada em margem (MS) combinada com KS no conjunto MNIST.

APÊNDICE B – FashionMNIST

B.1 Estratégias Individuais — Acurácia

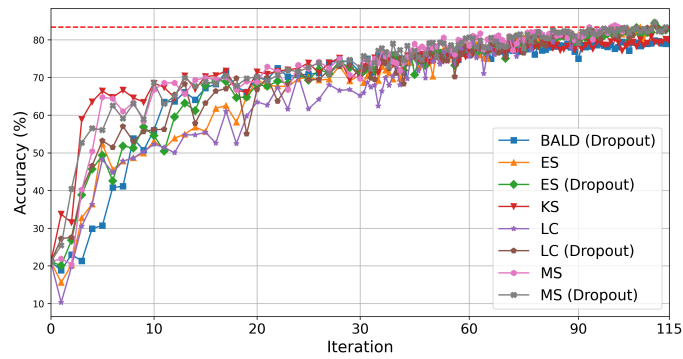
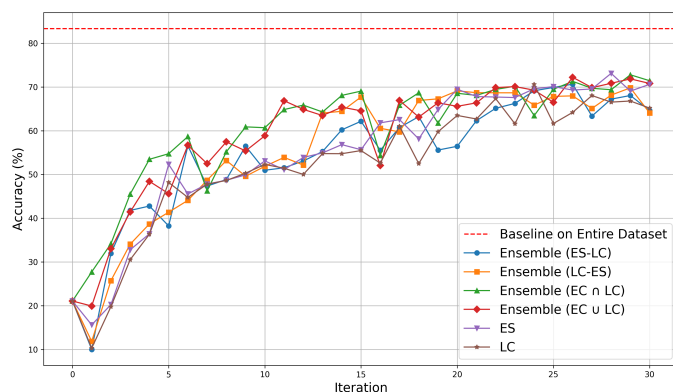
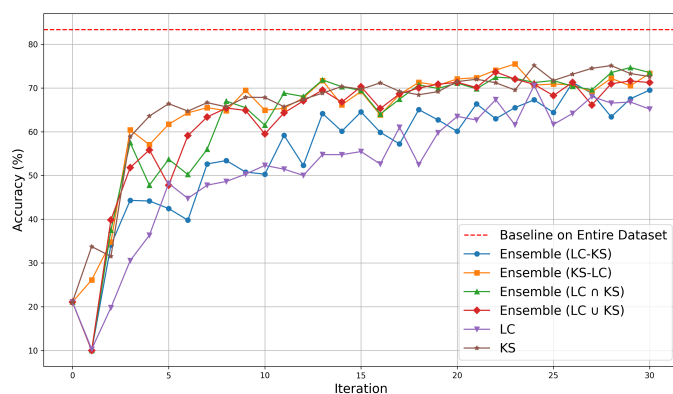


Figura 24 – Acurácia obtida ao executar cada estratégia até atingir 150 iterações ou a acurácia considerando todo o conjunto de dados mais três rounds para o conjunto FashionMNIST.

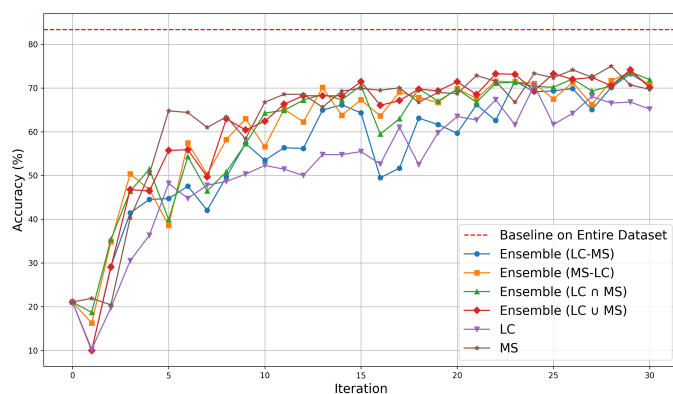
B.2 Combinações de Estratégias — Acurácia



(a)

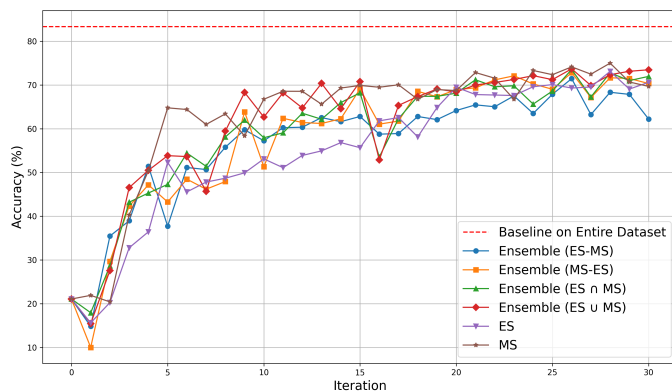


(b)

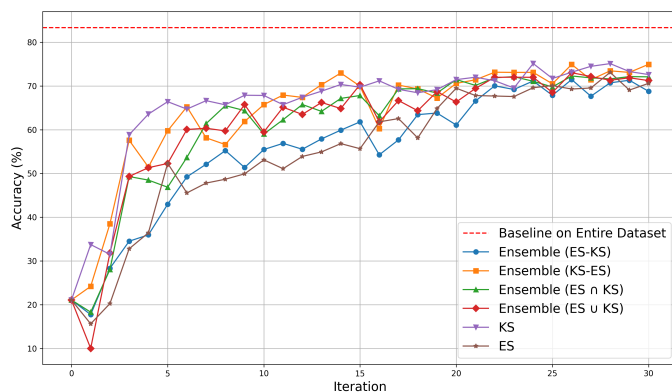


(c)

Figura 25 – Resultados de acurácia obtidos pela combinação da estratégia de aprendizado ativo baseada em incerteza (LC) com as estratégias individuais: (a) ES, (b) KS e (c) MS, respectivamente, utilizando o conjunto de dados FashionMNIST.



(a)



(b)

Figura 26 – Resultados de acurácia obtidos pela combinação da estratégia de aprendizado ativo baseada em entropia (ES) com as estratégias individuais: (a) MS e (b) KS, respectivamente, utilizando o conjunto de dados FashionMNIST.

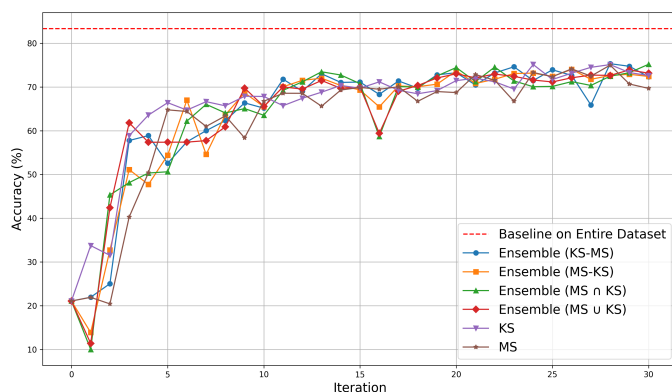
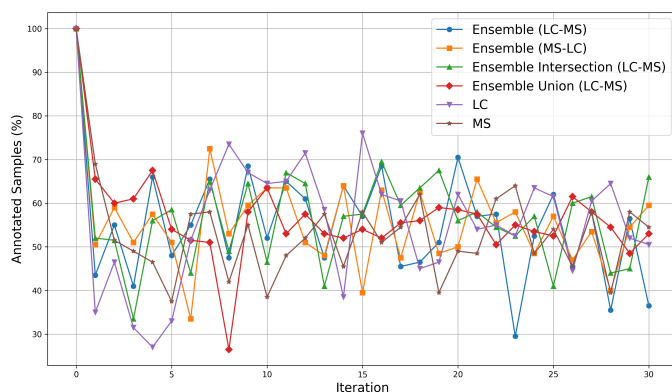
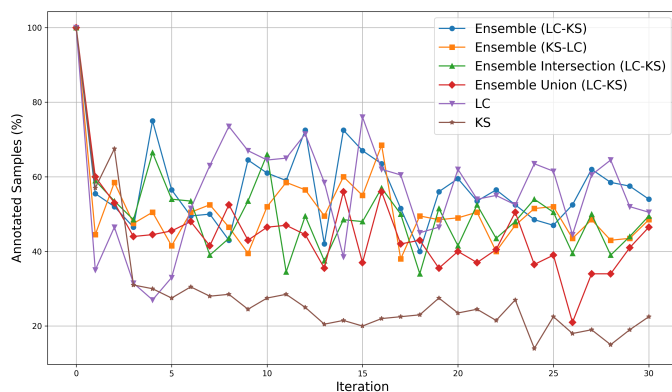


Figura 27 – Imagem do ensemble obtido pela combinação da estratégia de aprendizado ativo baseada em margem (MS) com KS, utilizando o conjunto de dados FashionMNIST.

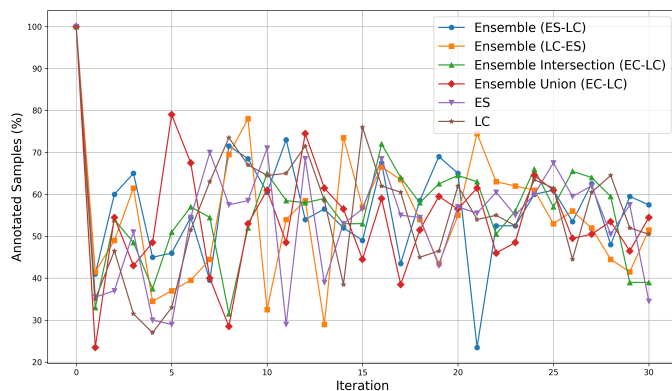
B.3 Combinações de Estratégias — Correções por Iterações



(a)

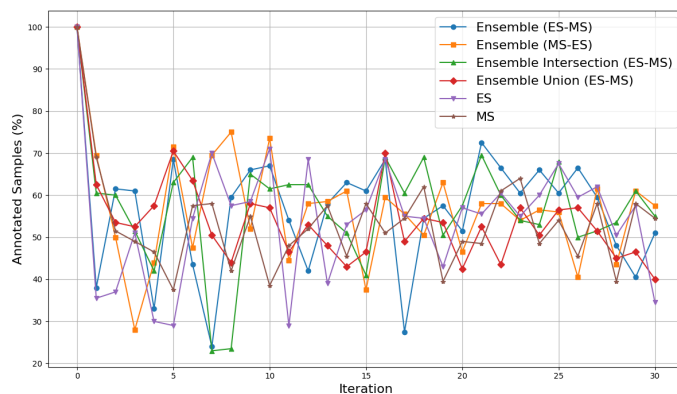


(b)

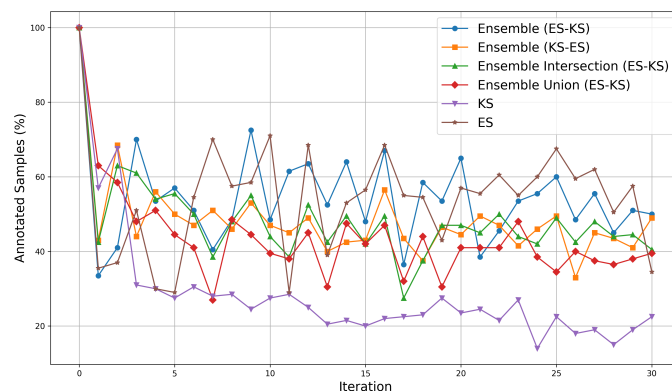


(c)

Figura 28 – Correções por iterações obtidas pela combinação da estratégia de aprendizado ativo baseada em incerteza (LC) com as estratégias individuais: (a) MS, (b) KS e (c) ES, respectivamente, utilizando o conjunto de dados FashionMNIST.



(a)



(b)

Figura 29 – Correções por iterações obtidas pela combinação da estratégia de aprendizado ativo baseada em entropia (ES) com as estratégias individuais: (a) MS e (b) KS, respectivamente, utilizando o conjunto de dados FashionMNIST.

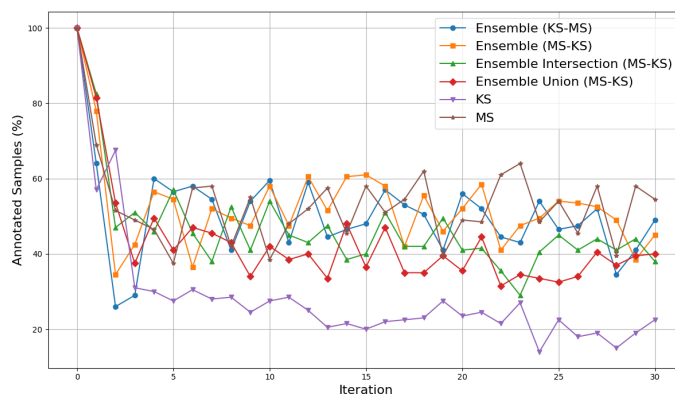
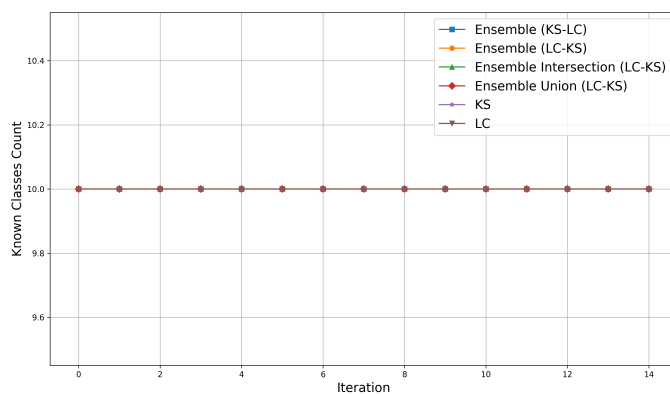
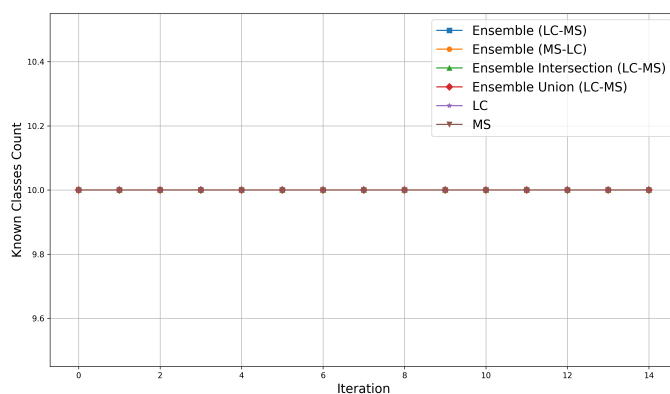


Figura 30 – Correção por iterações obtida pela combinação da estratégia de aprendizado ativo baseada em margem (MS) com KS, utilizando o conjunto de dados FashionMNIST.

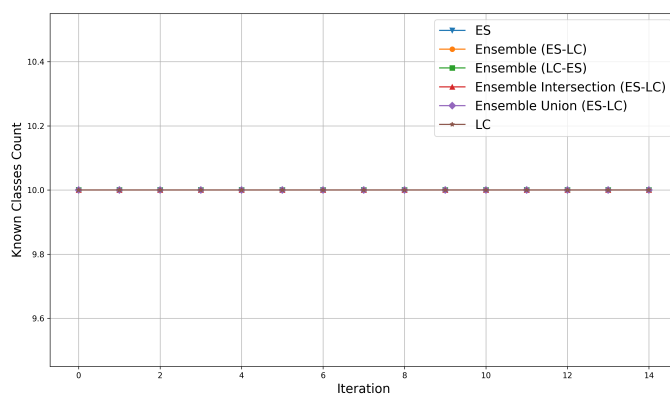
B.4 Combinações de Estratégias — Classes Conhecidas



(a)

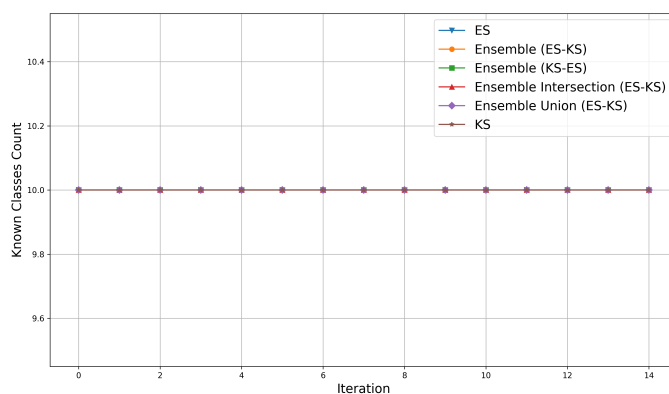


(b)

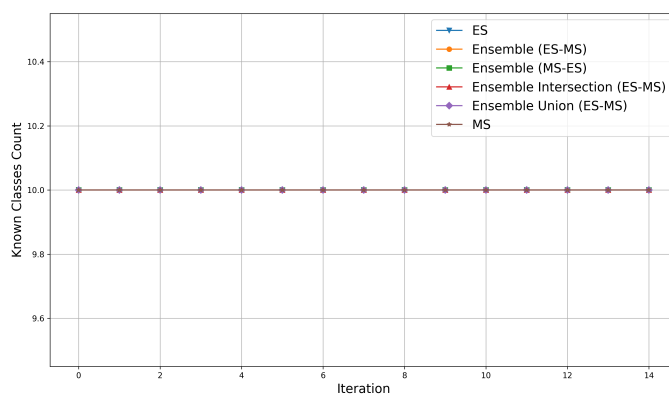


(c)

Figura 31 – Classes conhecidas ao longo das iterações utilizando a estratégia de aprendizado ativo baseada em incerteza (LC) combinada com: (a) KS, (b) MS e (c) ES, respectivamente, no conjunto de dados FashionMNIST.



(a)



(b)

Figura 32 – Classes conhecidas ao longo das iterações utilizando a estratégia de aprendizado ativo baseada em entropia (ES) combinada com: (a) KS e (b) MS, respectivamente, no conjunto de dados FashionMNIST.

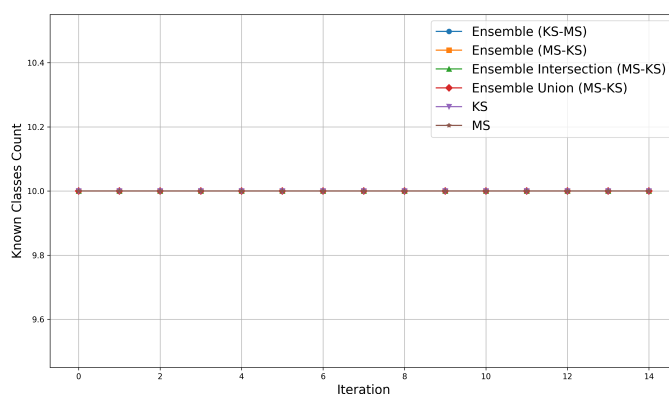


Figura 33 – Classes conhecidas ao longo das iterações utilizando a estratégia de aprendizado ativo baseada em margem (MS) combinada com KS, no conjunto de dados FashionMNIST.

APÊNDICE C – Parasitos

C.1 Estratégias Individuais — Acurácia

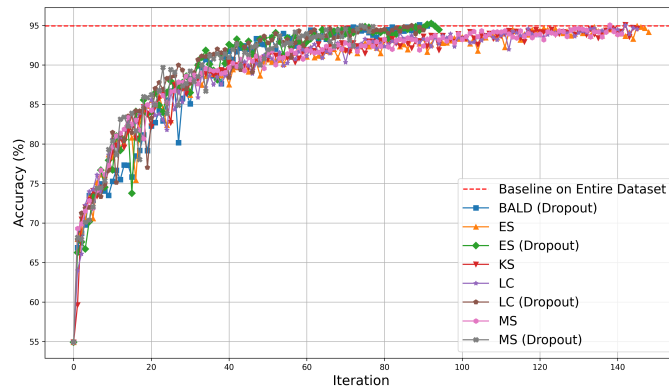
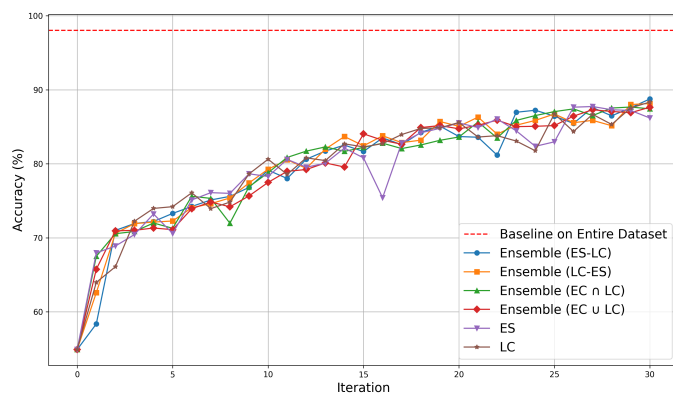
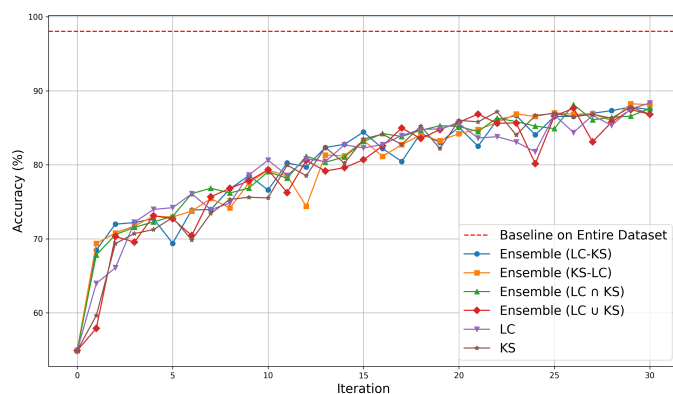


Figura 34 – Acurácia obtida ao executar cada estratégia até atingir 150 iterações ou a acurácia considerando todo o conjunto de dados mais três rounds para o conjunto de Parasitos.

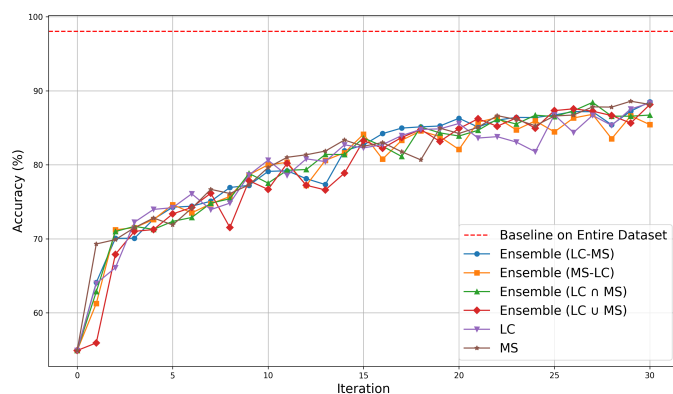
C.2 Combinações de Estratégias — Acurácia



(a)

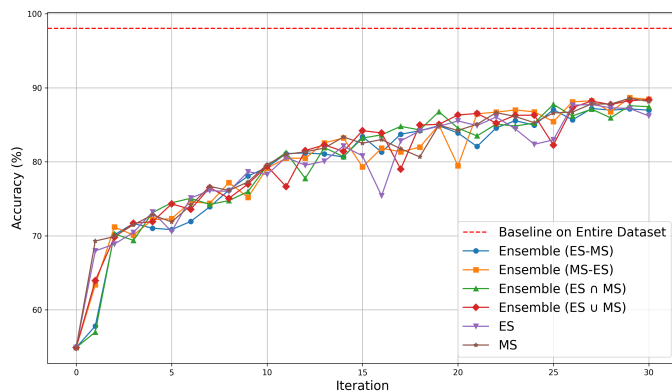


(b)

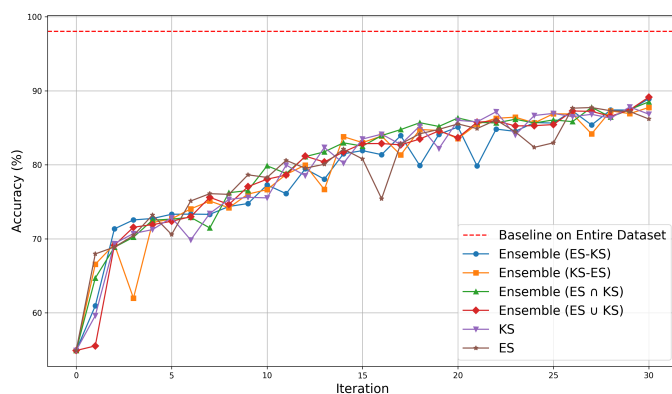


(c)

Figura 35 – Resultados de acurácia obtidos pela combinação da estratégia de aprendizado ativo baseada em incerteza (LC) com as estratégias individuais: (a) ES, (b) KS e (c) MS, respectivamente, utilizando o conjunto de dados Parasitos.



(a)



(b)

Figura 36 – Resultados de acurácia obtidos pela combinação da estratégia de aprendizado ativo baseada em entropia (ES) com as estratégias individuais: (a) MS e (b) KS, respectivamente, utilizando o conjunto de dados Parasitos.

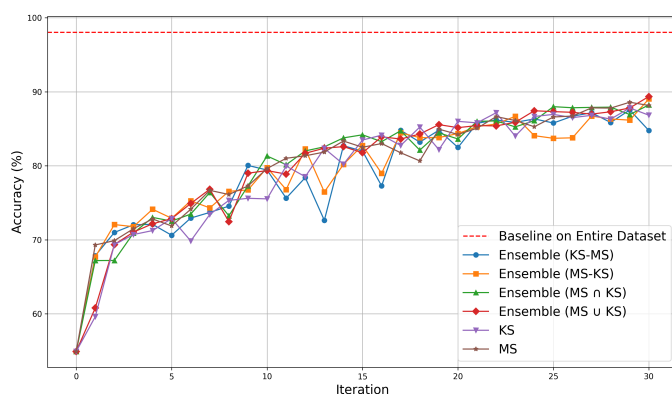
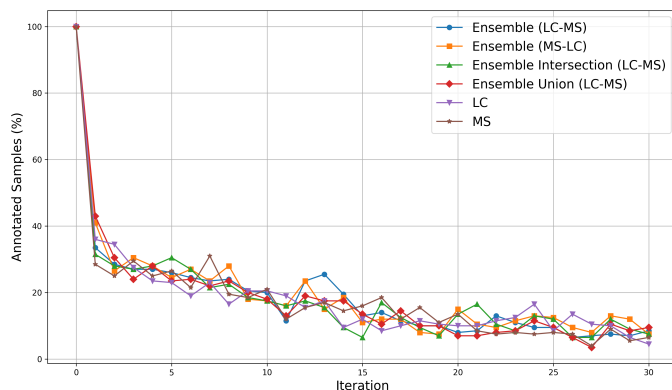
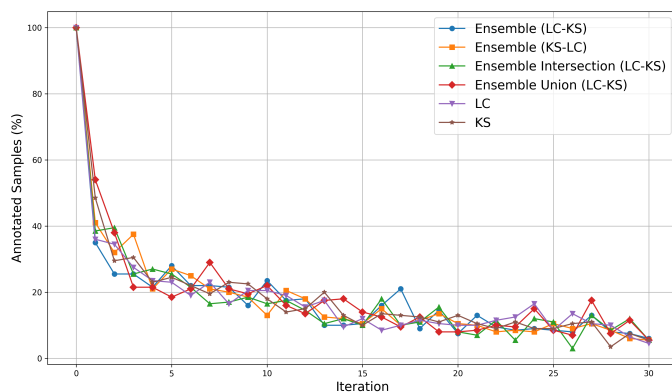


Figura 37 – Resultados de acurácia obtidos pela combinação da estratégia de aprendizado ativo baseada em margem (MS) com KS, utilizando o conjunto de dados Parasitos.

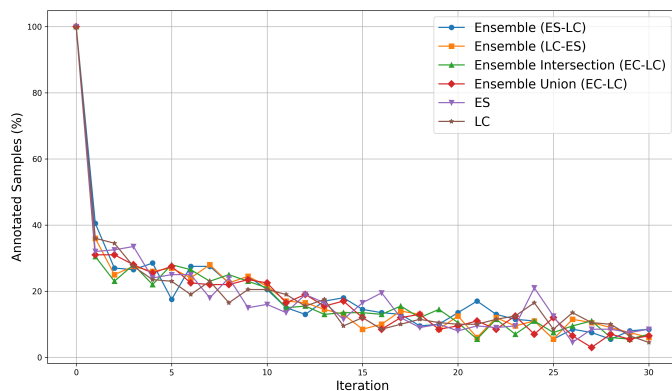
C.3 Combinações de Estratégias — Correções por Iterações



(a)

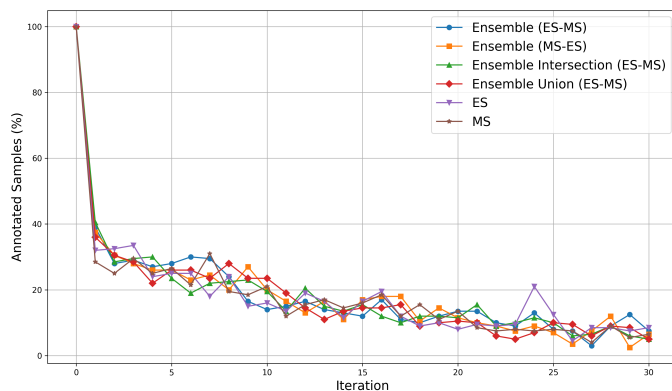


(b)

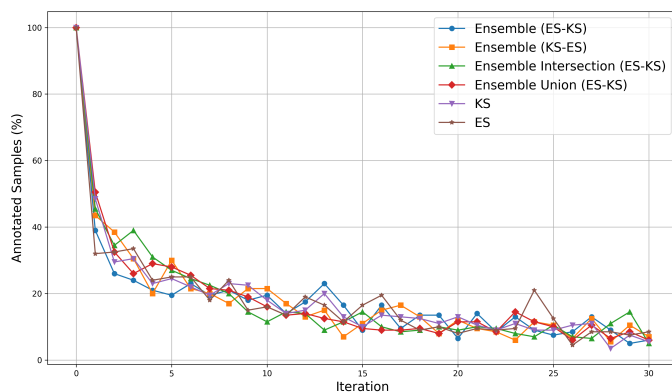


(c)

Figura 38 – Correções por iterações obtidas pela combinação da estratégia de aprendizado ativo baseada em incerteza (LC) com as estratégias individuais: (a) MS, (b) KS e (c) ES, respectivamente, utilizando o conjunto de dados Parasitos.



(a)



(b)

Figura 39 – Correções por iterações obtidas pela combinação da estratégia de aprendizado ativo baseada em entropia (ES) com as estratégias individuais: (a) MS e (b) KS, respectivamente, utilizando o conjunto de dados Parasitos.

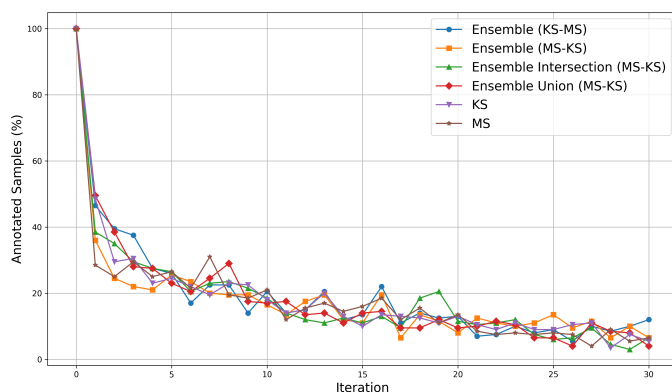
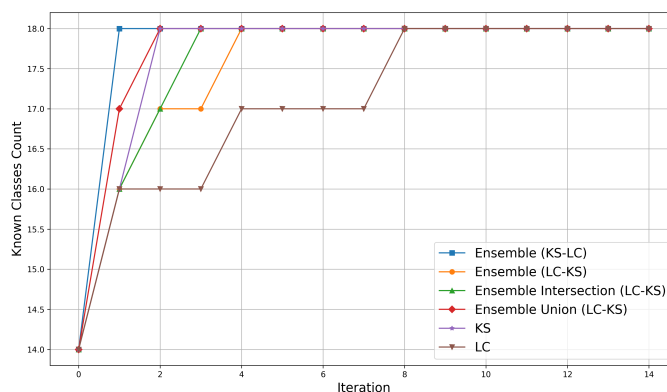
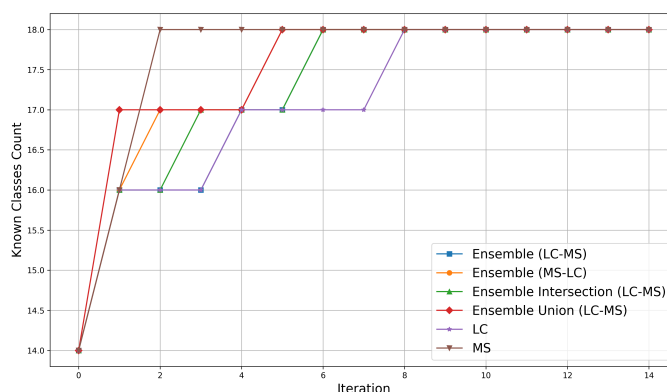


Figura 40 – Correção por iterações obtida pela combinação da estratégia de aprendizado ativo baseada em margem (MS) com KS, utilizando o conjunto de dados Parasitos.

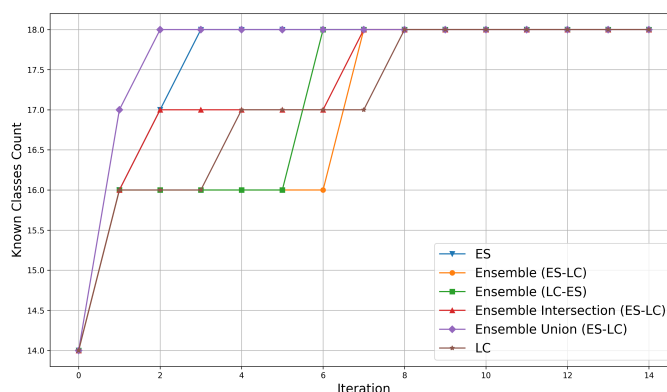
C.4 Combinações de Estratégias — Classes Conhecidas



(a)

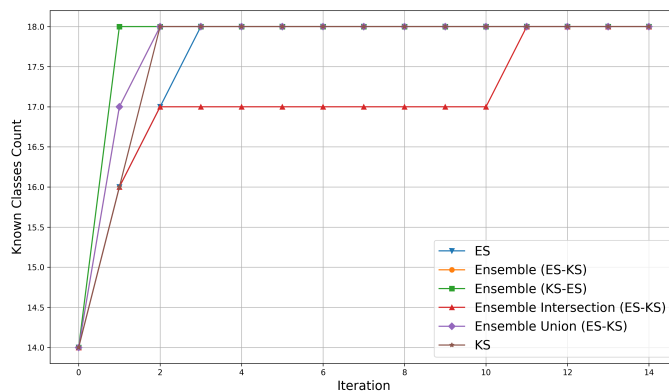


(b)

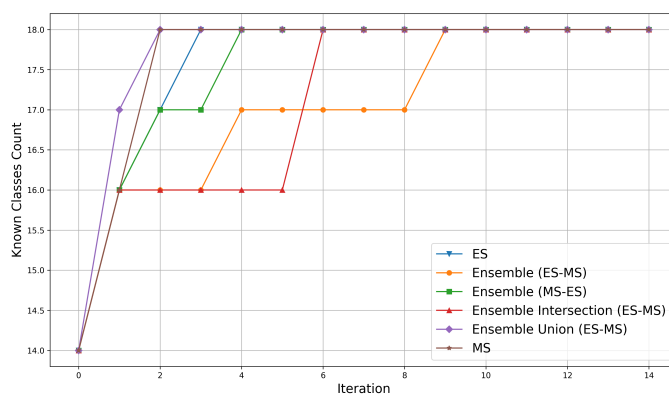


(c)

Figura 41 – Classes conhecidas ao longo das iterações utilizando a estratégia de aprendizado ativo baseada em incerteza (LC) combinada com: (a) KS, (b) MS e (c) ES, respectivamente, no conjunto de dados Parasitos.



(a)



(b)

Figura 42 – Classes conhecidas ao longo das iterações utilizando a estratégia de aprendizado ativo baseada em entropia (ES) combinada com: (a) KS e (b) MS, respectivamente, no conjunto de dados Parasitos.

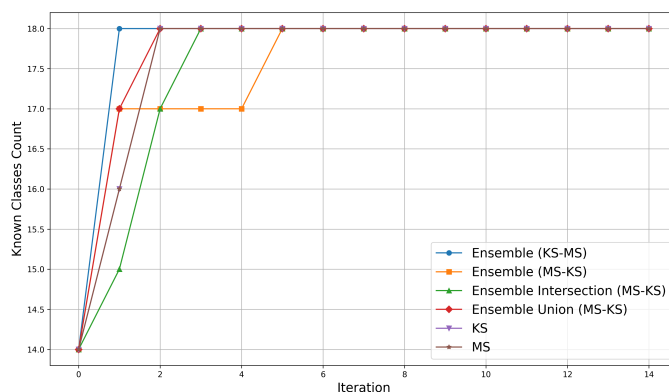


Figura 43 – Classes conhecidas ao longo das iterações utilizando a estratégia de aprendizado ativo baseada em margem (MS) combinada com KS, no conjunto de dados Parasitos.