Aprendizado Auto-Supervisionado para Reconhecimento de Atividade Humana baseado em Sensores

Palavras-Chave: ***Human activity recognition*****,** ***Self-Supervised*****,** **Contrastive-Learning**

Autores/as:

Pedro Hori Bueno [Instituto de Computação - Universidade Estadual de Campinas]

Profa. Dra. Sandra Eliza Fontes de Avila (orientadora) [Instituto de Computação - Universidade Estadual de Campinas]

Dra. Helena de Almeida Maia (coorientadora) [Instituto de Computação - Universidade Estadual de Campinas]

INTRODUÇÃO:

O aprendizado auto-supervisionado consiste, em sua essência, em fazer tarefas de predição, normalmente relacionadas ao aprendizado supervisionado, utilizando dados originalmente não rotulados e aplicando uma rotulação feita semi-automaticamente a partir dos próprios dados. Ele tem sido muito explorado na literatura devido às grandes vantagens que ele traz em relação à outras técnicas de aprendizado, como a não necessidade de supervisão humana na rotulação de dados e a abundância de dados não rotulados de fácil acesso.

Este trabalho foi desenvolvido dentro de um grupo de pesquisa que, juntamente ao instituto Eldorado, visa estudar arquiteturas cognitivas no contexto de Processamento de Linguagem Natural, com o intuito de auxiliar no desenvolvimento de uma tecnologia embarcada para dispositivos computacionais móveis (*smartphones*) que utiliza de tais arquiteturas para melhorar a experiência do usuário.

Desta forma, dialogando com o objetivo geral do grupo, este trabalho teve como foco o estudo de técnicas de aprendizado auto-supervisionado (do inglês, *Self-Supervised Learning*) em tarefas de Reconhecimento de Atividade Humanas (do inglês, *Human Action Recognition* - HAR) baseadas em sensores presentes em *smartphones* e em como estas técnicas performam em diferentes condições e bases de dados.

METODOLOGIA:

Neste projeto de iniciação científica foi feito um estudo de linguagem natural com o intuito de auxiliar no desenvolvimento de uma aplicação para arquiteturas móveis.

Primeiramente, foi feito um trabalho de pesquisa e levantamento bibliográfico na área de Processamento de Linguagem Natural e reconhecimento de atividades humanas, focando-se principalmente em trabalhos que utilizavam técnicas de aprendizado auto-supervisionado. Em seguida, foi feito o estudo e criação de códigos executáveis envolvendo os temas pesquisados anteriormente, assim como o estudo de bases de dados já existentes na literatura.

Durante todo o processo de pesquisa e desenvolvimento ocorreram reuniões entre os membros do projeto com o intuito de compartilhar conhecimento e definir as atividades do projeto.

Para um melhor entendimento do que foi desenvolvido no projeto de iniciação científica, segue abaixo os tópicos abordados:

- Pesquisa e levantamento bibliográfico em aprendizado auto-supervisionado;
- Pesquisa e levantamento bibliográfico de diferentes bases de dados para *HAR* na literatura;
- Estudo e desenvolvimento de códigos executáveis.

Aprendizado Auto-supervisionado

De acordo com [2] o termo “aprendizado auto-supervisionado" foi primeiramente introduzido na área de robótica, onde a rotulação dos dados é feita de maneira automatizada a partir da relação entre diferentes sinais de sensores. Posteriormente, o termo foi adaptado para o estudo de aprendizado de máquina e foi utilizado para se referir ao processo automatizado de geração de dados rotulados a partir de uma base de dados não rotulada, como descrito em [2].

O aprendizado auto-supervisionado pode ser visto como um subgrupo do paradigma de aprendizado não supervisionado, dado que ele não necessita de nenhum processo manual de rotulação dos dados. Porém, ao contrário do aprendizado não supervisionado que se baseia principalmente em identificação de padrões, agrupamentos ou detecção de anomalias nos dados, o aprendizado auto-supervisionado é utilizado em tarefas de predição, i.e., tarefas de aprendizado supervisionado.

Inicialmente, no projeto de iniciação científica, foi feita uma pesquisa sobre o que consiste a técnica de aprendizado auto-supervisionado. Nesta pesquisa, foram utilizados diversos sites de divulgação científica, assim como alguns artigos que abordam o tema de maneira didática, e a partir dela foi feita uma apresentação para o grupo de pesquisa sobre o tema. Foi estudado o aprendizado auto-supervisionado aplicado no contexto de imagens, vídeos e principalmente de *NLP*. Alguns resultados interessantes obtidos foram o melhor entendimento do tema de maneira ampla e um conhecimento de como esta área evoluiu ao longo do tempo com o surgimento de novas tecnologias.

Em seguida, foram estudados artigos científicos na área que abordavam diferentes meios de se aplicar o aprendizado auto-supervisionado. Em [7] foi estudado o aprendizado de diálogos a partir de uma rede neural que identifica a ordem de frases em um diálogo. Em [6], foi estudada a extração de contexto de um texto utilizando técnicas de pré-treinamento baseadas em tarefas de *mask*, *replace* e *switch*.

Dentro do campo de estudo de aprendizado auto-supervisionado, foi estudado com mais profundidade *Contrastive Learning* [1], uma abordagem discriminativa que tem como objetivo fazer o treinamento auto-supervisionado a partir do agrupamento de exemplos similares entre si e segregando exemplos distintos. Para isso, é determinada uma métrica de similaridade que mede quão próximos dois exemplos são, e o treinamento é feito classificando esta proximidade. Em *NLP*, esta abordagem é majoritariamente utilizada em combinação com técnicas de *Data Augmentation*, de modo que são aplicadas transformações como adição de ruídos ou inversão do sinal das amostras, e o treinamento auto-supervisionado consiste na detecção de similaridades entre as amostras alteradas com suas respectivas amostras originais.

Bases de Dados para HAR

Durante o projeto de iniciação científica foram estudadas com mais profundidade duas bases de dados no contexto de HAR, sendo elas a base de dados *KuHAR* [4] e a base de dados *MotionSense* [3]. Foram analisados aspectos relacionados aos processos de coleta, processamento e anotação de dados realizados pelas bases.

A base de dados *KuHAR* [4] foi criada para o uso de sensores presentes em *smartphones*, e outros dispositivos similares, para a resolução de tarefas de *HAR*. Nesta base, foram coletadas

amostras de 90 participantes, sendo 75 homens e 15 mulheres, de 18 a 34 anos, fazendo 18 atividades. Ao todo, foram coletadas 1945 amostras de atividades contendo dados de sensores acelerômetro e giroscópio a uma frequência de 100Hz.

A base de dados *MotionSense* [3] foi criada para o reconhecimento de atividades humanas a partir de sensores de movimento. A base de dados foi coletada a partir de 24 participantes, sendo 14 homens e 10 mulheres, fazendo 6 atividades distintas. Foram obtidos os valores de sensores acelerômetro e giroscópio, assim como valores de altitude e gravidade. Também foi anotado o gênero, idade, peso e altura de cada participante.

Estas bases de dados também foram utilizadas para o estudo e desenvolvimento de códigos executáveis para comparar a performance de diferentes técnicas de aprendizado auto-supervisionado através de bases de dados com características distintas.

Estudo e Desenvolvimento de Códigos Executáveis

Em paralelo ao desenvolvimento do trabalho e ao levantamento bibliográfico feito em busca da compreensão da literatura atual, também ocorreu uma procura mais específica por trabalhos que apresentassem códigos executáveis referentes aos artigos, possibilitando assim a realização de experimentos envolvendo a implementação e aprimoramento destes códigos.

Para isso, estudou-se o trabalho proposto em [5] que apresenta uma abordagem auto-supervisionada utilizando *contrastive learning* para solucionar um problema de classificação de ações baseadas em sensores. Este trabalho foi utilizado para a realização de diversos experimentos, pois seu código se encontrava disponível e funcional.

RESULTADOS E DISCUSSÃO:

Primeiramente, foi feito um trabalho de compreensão e implementação do código apresentado em [5]. O mesmo foi executado por membros da equipe e os resultados obtidos foram comparados com os resultados apresentados no artigo original. A base de dados utilizada pelo artigo para o estudo do código foi a *MotionSense*.

O código foi facilmente executado e os resultados obtidos, ainda que diferentes, foram muito similares aos resultados apresentados no artigo.

No artigo foi utilizado a métrica F1, que corresponde à média harmônica entre a precisão e o *recall*, para avaliar o resultado obtido. O artigo obteve um valor da métrica F1 igual a 0,944, enquanto na reprodução do código o valor da métrica F1 obtido foi de 0,889.

Atribui-se essa diferença de resultados a variáveis aleatórias presentes no código que, por sua vez, são dificilmente contornáveis.

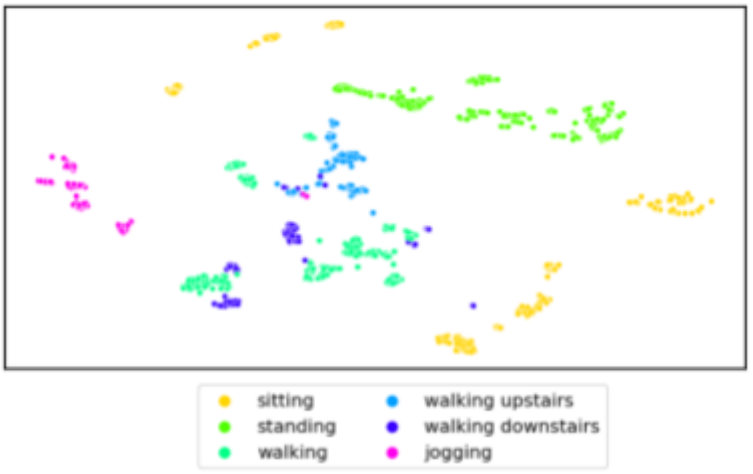

Imagem 1 - Representação gráfica do resultado obtido na base de dados MotionSense

Em seguida, houve um trabalho de adaptação e aprimoramento do código. Esta etapa teve como objetivo estudar a aplicabilidade do código para diferentes bases de dados, e como a estratégia apresentada no artigo se comporta em um contexto mais amplo. A *MotionSense*, originalmente usada, é uma base de dados relativamente simples, possuindo um número baixo de atividades e amostras. Desta forma, tentou-se adaptar o código para o uso de outra base de dados mais completa, com um

número superior de atividades e amostras. Assim, foi escolhida a base *KuHAR*, e o código foi adaptado para este novo conjunto de dados.

As principais alterações feitas no código foram a implementação dos dados de um novo sensor, o giroscópio, que não era utilizado no código original mas estava presente e era muito relevante na base de dados *KuHAR*, e a alteração da leitura dos dados e sua separação em conjuntos de treinamento, validação e teste.

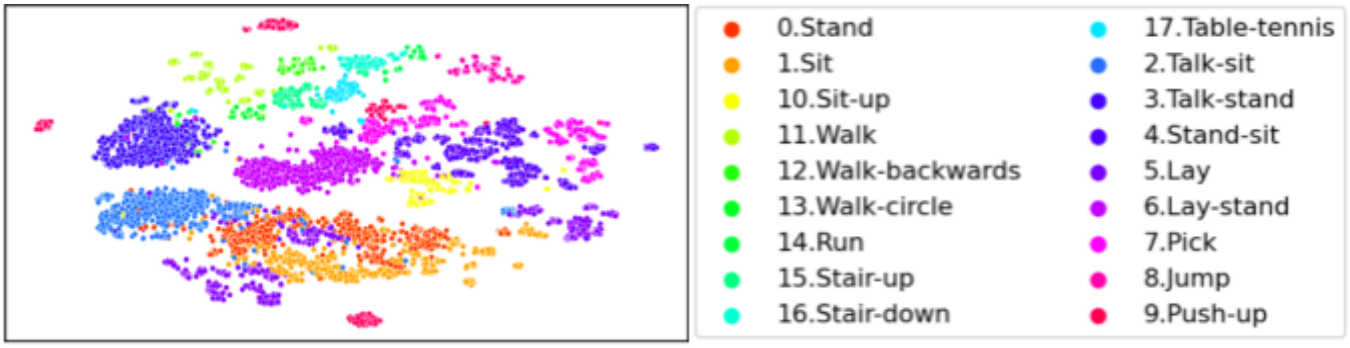

Imagem 2 - Representação gráfica do resultado obtido na base de dados KuHAR

O resultado obtido na base de dados *KuHAR* completa utilizando-se a métrica F1 foi de 0,905 (tabela 1), superior ao resultado de 0,876 apresentado no artigo original que propôs a base, porém inferior aos resultados obtidos pelo atual estado da arte referente à esta base.

	Fine-tunning: Completo	Fine-tunning: Acelerômetro	Fine-tunning: Giroscópio
Pre-treinamento: Completo	0,905 ±0,008	0,86 ±0,01	0,81 ±0,02
Pre-treinamento: Acelerômetro	0,88 ±0,01	0,86 ±0,01	0,76 ±0,02
Pre-treinamento: Giroscópio	0,83 ±0,03	0,86 ±0,01	0,81 ±0,01
Pre-treinamento: Ausente	0,77 ±0,03	0,74 ±0,03	0,68 ±0,02

Tabela 1 - Diferentes dados de Pré-Treinamento e Fine-Tunning na base KuHAR

Também foi adicionado ao código um processo de treinamento cruzado, em que o pré-treinamento da rede neural era feito em um conjunto de dados contendo determinados sensores e o *Fine-Tuning* da rede era realizado em um conjunto de dados contendo uma combinação diferente de sensores (tabela 1). Tais experimentos são importantes para analisar como a presença de cada sensor contribui para um melhor desempenho do código. Também foram realizados experimentos em que o pré-treinamento não era realizado e a rede era inicializada de maneira aleatória, mostrando assim como o pré-treinamento influencia no desempenho.

Por fim, foi feito um cruzamento entre diferentes bases de dados de modo que a rede neural realizasse seu pré-treinamento em uma base de dados, e o processo de *Fine-Tuning* em outra, sendo elas as bases *KuHAR* e *MotionSense* (Tabela 2). Concluímos assim que ao realizar o pré-treinamento em uma base de dados mais simples o desempenho do algoritmo cai significativamente para um

	Fine-tunning: Acelerômetro KuHAR	Fine-tunning: Acelerômetro MotionSense
Pre-treinamento: Acelerômetro KuHAR	0,86 ±0,01	0,91 ±0,01
Pre-treinamento: Acelerômetro MotionSense	0,65 ±0,03	0.93 ±0,01

Tabela 2 - Treinamento Cruzado entre diferentes bases de dados

Fine-Tuning em uma base mais complexa, porém, ao realizar o pré-treinamento em uma base de dados mais complexa, este pré-treinamento pode ser utilizado para o *Fine-Tuning* de uma base mais simples sem uma perda significativa de performance.

Futuramente, pretende-se aprimorar ainda mais o código, de modo que seja possível comparar os resultados de classificação obtidos por ele com os resultados por outras técnicas de aprendizado auto-supervisionado e algumas técnicas comuns de aprendizado supervisionado utilizadas na literatura.

CONCLUSÕES:

Durante este projeto de iniciação científica, foi possível realizar uma ampla pesquisa no tópico de aprendizado auto-supervisionado e um levantamento de diferentes bases de dados utilizadas em HAR disponíveis na literatura. Assim, contribuiu-se para a produção de 2 relatórios técnicos entregues pela Meta 5 do Hub de Inteligência Artificial e Arquiteturas Cognitivas H.IAAC.

Também foi estudado e desenvolvido um código executável que utiliza técnicas de aprendizado auto-supervisionado para tarefas de HAR, mostrando o desempenho desta técnica em diferentes situações e possibilitando um melhor entendimento da área.

BIBLIOGRAFIA

[1] Ashish Jaiswal, Ashwin Ramesh Babu, Mohammad Zaki Zadeh, Debapriya Banerjee, e Fillia Makedon. A Survey on Contrastive Self-supervised Learning. Technologies, 2020.

[2] Xiao Liu, Fanjin Zhang, Zhenyu Hou, Li Mian, Zhaoyu Wang, Jing Zhang, e Jie Tang. Self- supervised Learning: Generative or Contrastive. IEEE Transactions on Knowledge and Data Engineering, 2021.

[3] Mohammad Malekzadeh, Richard G. Clegg, Andrea Cavallaro, e Hamed Haddadi. Protecting Sensory Data Against Sensitive Inferences. Em 1st Workshop on Privacy by Design in Distributed Systems, 2018.

[4] Niloy Sikder e Abdullah-Al Nahid. KU-HAR: An Open Dataset for Heterogeneous Human Activity Recognition. Pattern Recognition Letters, 2021.

[5] Chi Ian Tang, Ignacio Perez-Pozuelo, Dimitris Spathis, e Cecilia Mascolo. Exploring Contrastive Learning in Human Activity Recognition for Healthcare. arXiv preprint arXiv:2011.11542, 2020.

[6] Hong Wang, Xin Wang, Wenhan Xiong, Mo Yu, Xiaoxiao Guo, Shiyu Chang, e William Yang Wang. Self-supervised learning for contextualized extractive summarization. arXiv preprint arXiv:1906.04466, 2019.

[7] Jiawei Wu, Xin Wang, e William Yang Wang. Self-supervised dialogue learning. arXiv preprint arXiv:1907.00448, 2019.