

Automation: Uma Ferramenta para a Evolução de um Tradutor Automático Português-Libras com Corpus Colaborativo

Marcos Henrique Alves da Silva¹, Tiago Maritan Ugulino de Araújo¹,
Rostand Edson Oliveira Costa¹, Samuel de Moura Moreira¹,
Lucas Moreira e Silva Alves¹

¹Centro de Informática – Universidade Federal da Paraíba (UFPB)
João Pessoa – Paraíba – Brasil

{marcos.alves, rostand, samuel.moura, tiagomaritan}@lavid.ufpb.br

lmsa.moreira@gmail.com

Abstract. *Deaf people communicate naturally using visuospatial languages, called Sign Languages (SL). Although Sign Languages are recognized in many countries as a second official language, content is not always generated in an accessible way for this community. In order to minimize these impacts, research involving automatic translation using Deep Learning is gaining more and more space in the literature. However, one of the main barriers to using this strategy is the need for a bilingual corpus, which is not always an easy task to find or produce. In this context, this paper presents a training automation tool for a Portuguese-Libras automatic translator. The solution communicates with a collaborative system of translations of Portuguese-Libras sentences, WikiLibras, and feeds the corpus of an automatic translator for Libras, VLibras, with the goal of conducting training with an increasingly robust data set each iteration.*

Resumo. *Pessoas surdas comunicam-se naturalmente usando línguas visoespaciais, denominadas Línguas de Sinais (LS). Apesar das línguas de sinais serem reconhecidas em muitos países como uma segunda língua oficial, nem sempre os conteúdos são gerados de forma acessível para esta comunidade. Com objetivo de minimizar esses impactos, pesquisas envolvendo tradução automática utilizando Aprendizagem Profunda ganham cada vez mais espaço na literatura. Porém, uma das principais barreiras para uso dessa estratégia é a necessidade de um corpus bilíngue, o que nem sempre é uma tarefa fácil de encontrar ou produzir. Nesse contexto, este trabalho apresenta uma ferramenta de automação de treinamentos para um tradutor automático Português-Libras. A solução comunica-se com um sistema colaborativo de traduções de sentenças Português-Libras, o WikiLibras, e alimenta o corpus de um tradutor automático para Libras, o VLibras, com objetivo de realizar treinamentos com um conjunto de dados cada vez mais robusto a cada iteração.*

1. Introdução

Segundo a *World Health Organization* (WHO)¹ mais de 5% da população mundial possui algum nível de deficiência auditiva, o que representa cerca de 466 milhões de pessoas. Segundo o estudo realizado pelo Instituto Locomotiva e a Semana da Acessibilidade Surda², no Brasil, cerca de 10,7 milhões de pessoas possuem algum tipo de deficiência auditiva. Desse total, 2,3 milhões têm deficiência severa.

Os surdos enfrentam uma série de barreiras que lhe são apresentadas desde os primeiros anos de vida, seja no ambiente familiar ou educacional, causando malefícios que perduram até mesmo na fase adulta [Rosa Zucolotto et al. 2019]. Eles tem dificuldade principalmente na comunicação com pessoas ouvintes e no acesso à informação, uma vez que as pessoas ouvintes comunicam-se cotidianamente utilizando as línguas orais (LO), enquanto eles, por sua vez, comunicam-se naturalmente por intermédio das línguas de sinais (LS).

Como uma alternativa prática e complementar para auxiliar o acesso à informação das pessoas surdas, na literatura científica, algumas ferramentas de tradução automática de conteúdos nas línguas orais para línguas de sinais vêm sendo desenvolvidas, como, por exemplo, os trabalhos de [Vichyaloetsiri et al. 2017] e [Krishna et al. 2020]. Elas geralmente tem como objetivo atender e diminuir os impactos causados pelo difícil acesso a LS para comunidade surda. Elas podem ser úteis principalmente em casos onde existe um grande volume e dinamicidade dos conteúdos, como, por exemplo, na *Web*, o que dificulta o atendimento a oferta de conteúdos em línguas de sinais usando apenas intérpretes humanos.

No Brasil, as ferramentas Hand Talk³ e VLibras [Araujo 2012] destacam-se na oferta de tradução automática de texto e áudio em Português para Língua Brasileira de Sinais (Libras), possibilitando o acesso de conteúdos na internet em Libras, bem como auxiliam no uso diário por meio de aplicações para dispositivos móveis. Tanto no caso do VLibras quanto do Handtalk, a sinalização em Libras é feita por meio de um intérprete animado, também chamado de avatar 3D [Veríssimo et al. 2019].

No entanto, existe uma série de desafios para o desenvolvimento desses tradutores automáticos. Um desses desafios é a dificuldade em se encontrar ou produzir corpus bilíngues envolvendo língua oral (ex: Português) e a língua de sinais (ex: Libras), que são essenciais para a aplicação de técnicas que estão no estado da arte em tradução automática. Nesse contexto, é importante mencionar que, atualmente, o estado da arte e da técnica em tradução automática entre línguas orais, por exemplo, baseia-se em técnicas de Aprendizagem Profunda (do inglês, *Deep Learning*), que geralmente demandam grandes bases de dados para que apresentem uma boa qualidade de tradução como apresentado no trabalho de [Liu et al. 2019].

Para endereçar esse problema e tentar gerar traduções automáticas para língua de sinais de melhor qualidade, neste trabalho é proposta uma abordagem incremental au-

¹<https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss>

²<https://agenciabrasil.ebc.com.br/geral/noticia/2019-10/brasil-tem-107-milhoes-de-deficientes-auditivos-diz-estudo>

³<https://www.handtalk.me/>

tomatizada (Automation) para o treinamento de um tradutor automático para Libras. A solução comunica-se com um sistema colaborativo de traduções de sentenças Português-Libras, o WikiLibras, e alimenta o corpus de um tradutor automático para Libras, o VLibras, com objetivo de realizar treinamentos com um conjunto de dados cada vez mais robusto a cada iteração.

O restante do artigo está organizado da seguinte forma. Na Seção 2, será apresentada a solução proposta, descrevendo cada um dos componentes; na Seção 3, os experimentos realizados para validar a solução; a Seção 4 apresenta uma análise dos resultados dos experimentos; por fim, a Seção 5 contém as considerações finais do trabalho.

2. Ferramenta Proposta

Nessa seção serão apresentados a arquitetura e os componentes da ferramenta proposta de maneira detalhada. A solução foi integrada com as ferramentas da Suíte VLibras. Essa suíte conta com um tradutor automático para Libras de código aberto, denominado VLibras⁴, produzido em parceria entre o Ministério da Economia (ME), por meio da Secretaria de Governo Digital (SGD) e a Universidade Federal da Paraíba (UFPB). Além disso, a suíte conta com um sistema denominado WikiLibras⁵, que tem como propósito que pessoas surdas, especialistas em Libras e colaboradores possam contribuir com a evolução do sistema sugerindo novos sinais, propondo correções nos sinais, bem como, sugerindo novos exemplos de tradução Português-Libras. Por fim, é possível destacar o VLibras Console, que é um sistema de gerenciamento, controle e operação das ferramentas e serviços da Suíte VLibras.

2.1. Visão Geral

Conforme mencionado na Seção 1, a solução tem como objetivo automatizar e gerenciar o processo de treinamento do tradutor automático VLibras de maneira incremental. Uma visão esquemática da solução integrada com a Suíte VLibras é apresentada na Figura 1. A ferramenta funciona da seguinte forma: Inicialmente, os usuários contribuem com propostas de novos sinais no sistema WikiLibras, assim que esses sinais são aprovados por um conjunto de especialistas, o componente Serviço de Integração, que utiliza o protocolo HTTP, realiza uma consulta e os envia automaticamente para o serviço Automation que por sua vez gerencia o corpus e o treinamento do tradutor automático VLibras por intermédio de chamadas de sistemas. Por fim, todas as ações de treinamento do tradutor podem ser realizadas pela interface do VLibras-Console que realiza chamadas HTTP para o componente Automation.

A Figura 2 apresenta a comunicação entre o WikiLibras, Serviço de Integração e o Serviço Automation quando um novo sinal é cadastrado por algum colaborador do WikiLibras. O Serviço de Integração sincroniza os novos sinais do WikiLibras de maneira automática e os envia para o Serviço Automation que retorna a informação indicando se os novos sinais foram corretamente inseridos.

A Figura 3 apresenta a comunicação entre o VLibras Console, Serviço Automation e o Tradutor VLibras quando o operador realiza alguma ação na ferramenta VLibras Console. O Serviço Automation recebe todas as ações (ex. iniciar treinamento, consultar

⁴<https://www.vlibras.gov.br>

⁵<https://wiki.vlibras.gov.br>

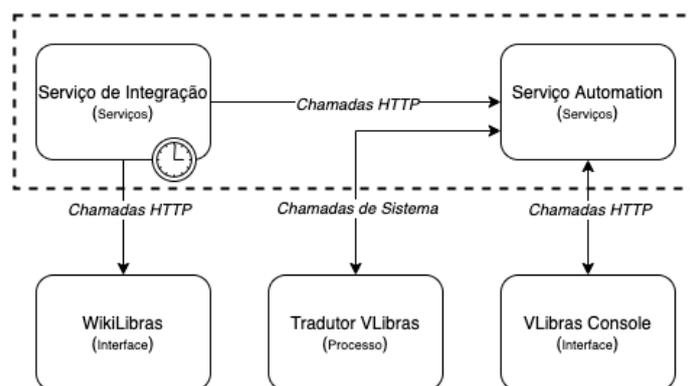


Figura 1. Visão esquemática da solução proposta integrada com a Suíte VLibras

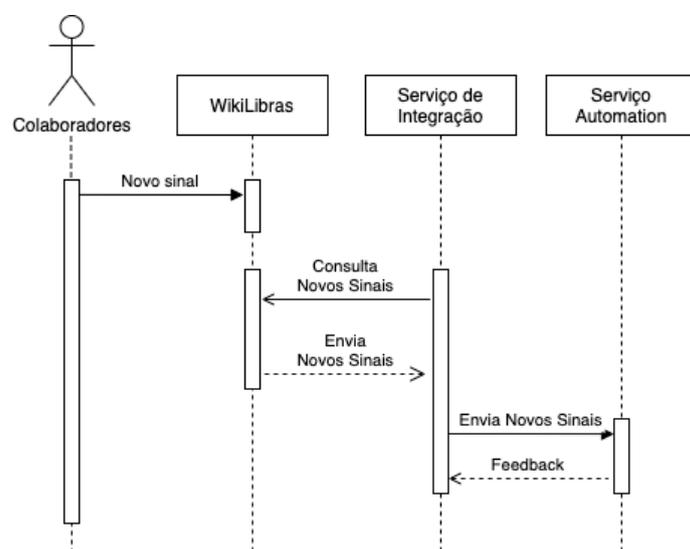


Figura 2. Diagrama da solução proposta integrada com WikiLibras

treinamento etc.) que o operador informa no VLibras Console, faz o tratamento e instancia o Tradutor VLibras para executar a ação requisitada.

Nas próximas seções serão apresentados os dois serviços que compõem a ferramenta Automation e sua integração com a Suíte VLibras.

2.2. Serviço de Integração

O componente de Integração é um processo configurável que acessa uma base de sentenças e encaminha diretamente para o serviço Automation. A Tabela 1, apresenta as principais variáveis do serviço.

O componente de Integração foi escrito em TypeScript utilizando Node.js, além disso foi utilizado o banco de dados Redis para memorização de todas as sentenças enviadas do WikiLibras para o componente Automation.

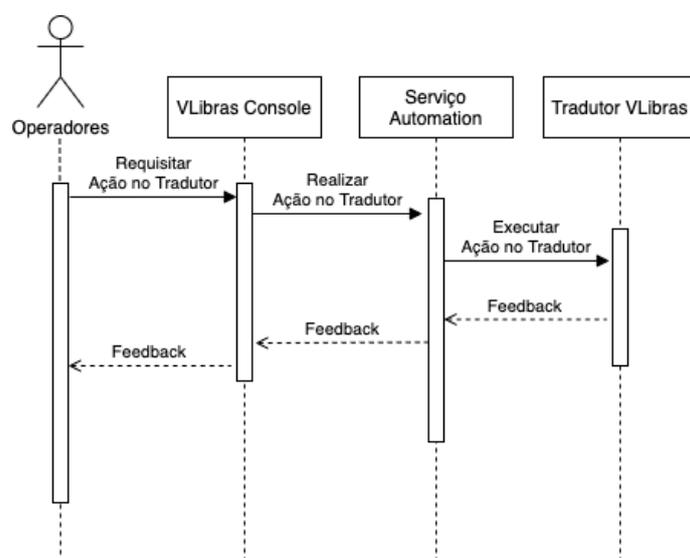


Figura 3. Diagrama da solução proposta integrada com VLibras Console e o Tradutor

Tabela 1. Parâmetros configuráveis do Serviço de Integração

Parâmetros	Descrição
URL SOURCE	Local que o serviço irá coletar as frases
URL TARGET	Local que o serviço irá enviar as frases
CLOCK	O espaço de tempo que o serviço deve realizar a coleta
MAX PHRASE	O máximo de frases que deve ser coletada a cada checagem

2.2.1. Integrando Serviço de Integração ao WikiLibras

Conforme mencionado na Seção 2, o WikiLibras é uma ferramenta colaborativa que tem como propósito coletar sugestões da comunidade surda para melhorias no seu vocabulário de sinais e na sua base de exemplos de tradução.

Neste trabalho, o componente Serviço de Integração foi integrado ao WikiLibras para enviar os exemplos de tradução aprovados pelos linguistas, especialistas em Libras, cadastrados no sistema, para o serviço Automation. Este último, por sua vez, faz o tratamento da sentença e envia diretamente para o corpus de treinamento do tradutor automático do VLibras. A Figura 4 apresenta a interface utilizada pelos especialistas para a avaliação de novos sinais submetidos ao WikiLibras.

A partir dessa modificação no WikiLibras será possível utilizar os exemplos de tradução Português-Libras produzidos pela comunidade surda e demais colaboradores, para posterior evolução do corpus de tradução do VLibras e, conseqüentemente, para potencial melhoria do seu componente de tradução automática para Libras.

Essas últimas ações, evolução do corpus de tradução e realização de novos treinamentos do componente de tradução automática do VLibras, serão realizadas pelo componente Automation que será integrado ao VLibras-Console conforme será apresentado na próxima seção.

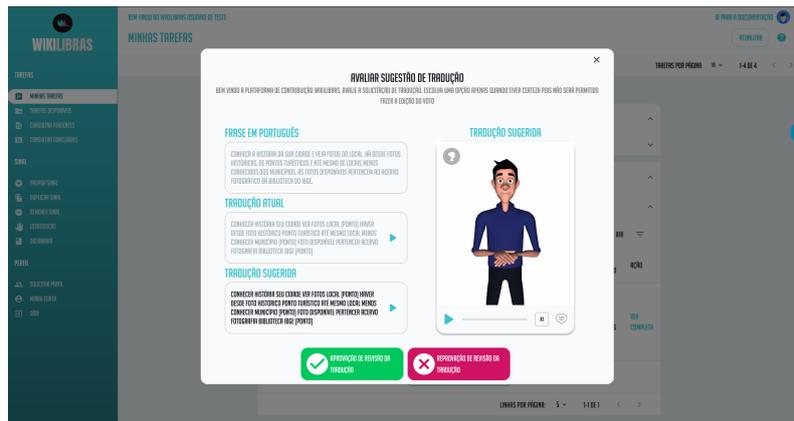


Figura 4. Interface de Revisão de Tradução do WikiLibras

2.3. Serviço Automation

O componente Automation é uma API que fornece uma interface HTTP permitindo a integração entre o tradutor automático VLibras e o sistema VLibras-Console. A Figura 5 apresenta a arquitetura do componente de Automation. A API é dividida em duas camadas. A camada de serviços é responsável por toda lógica de negócio da aplicação, cada serviço acessa o banco por intermédio das entidades, que por sua vez possuem acesso direto a base de dados.

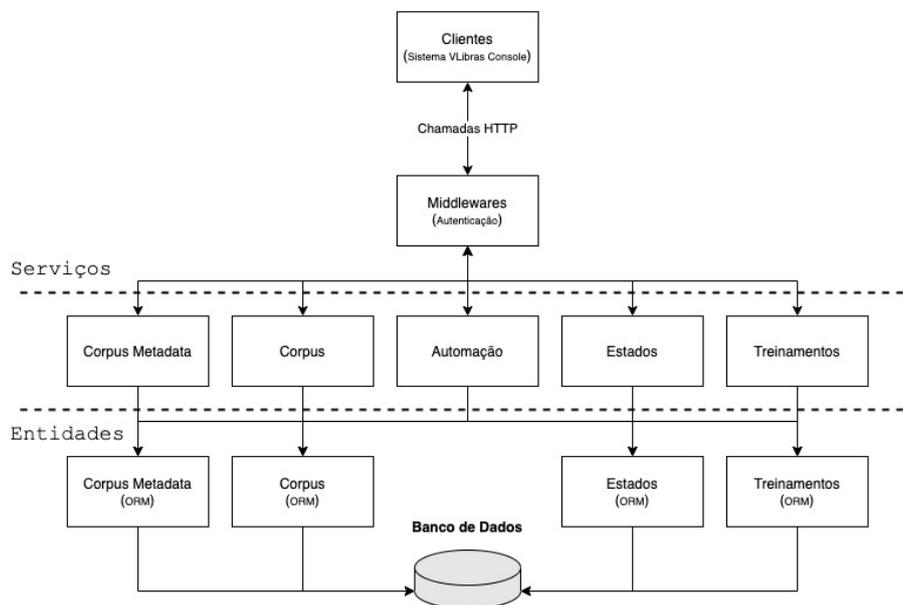


Figura 5. Arquitetura do componente Automation

De acordo com a Figura 5, cada classe de serviço possui as seguintes funções no sistema:

1. Corpus Metadata - Tem como objetivo gerenciar cada versão do corpus;
2. Corpus - Gerencia todas as sentenças que serão utilizadas no treinamento. Cada sentença possui um versão associada, o que possibilita o rastreamento de quais sentenças foram utilizadas em cada treinamento;

3. Estados - Gerencia o estado do treinamento do tradutor automático. Os estados atualmente suportados são: Pendente, Progresso, Finalizado, Falhou e Cancelado;
 - (a) Pendente - Treinamento está apenas instanciado mas ainda não foi iniciado;
 - (b) Progresso - Treinamento está em processo de execução;
 - (c) Finalizado - Treinamento foi encerrado de maneira bem-sucedida;
 - (d) Falhou - Treinamento falhou por alguma razão;
 - (e) Cancelado - Treinamento foi cancelado pela ação do usuário;
4. Treinamentos - Gerencia todas as informações necessárias para instância de um treinamento;
5. Automação - Apresenta funções, que podem ser acessadas pelo usuário por intermédio da ferramenta VLibras-Console, com a finalidade de gerenciar um tradutor automático;

O componente Automation foi escrito em Python utilizando o framework Falcon que fornece uma interface para construção de APIs HTTP integrado com banco de dados PostgreSQL.

2.3.1. Integrando Automation ao Tradutor Automático VLibras

O *pipeline* do tradutor automático VLibras utiliza a ferramenta *Facebook AI Research Sequence-to-Sequence Toolkit* (Fairseq) [Ott et al. 2019]. O Fairseq é uma ferramenta que possibilita a modelagem em alto nível para diferentes aplicações (e.g. tradução e sumarização).

O componente Automation comunica-se com Tradutor VLibras por intermédio de chamadas de sistema que criam uma instância de execução do *pipeline*. Para cada chamada, o *pipeline* devolve uma resposta equivalente.

2.3.2. Integrando Automation ao VLibras Console

Por fim, o serviço Automation também foi integrado ao VLibras-Console, para permitir realização de novos treinamentos no componente de tradução automática para Libras, a partir dos exemplos aprovados pelos especialistas no WikiLibras. Conforme mencionado na Seção 2, o VLibras Console é um sistema de gerenciamento, controle e operação de todas as aplicações da Suíte VLibras.

No processo de integração do VLibras-Console com o serviço Automation, foram implementadas novas telas de interface dentro do VLibras-Console para permitir que o operador execute as funções previstas no componente Automation.

A Figura 6 apresenta uma das interfaces adicionadas para dar suporte a visualização de treinamentos do tradutor VLibras por intermédio do componente Automation. O operador possui acesso a evolução do treinamento, bem como as ações de iniciar, parar e testar um modelo.

3. Experimentos

Nessa seção será apresentada o experimento conduzido com o objetivo de avaliar a solução proposta. Para isso, foi projetado um experimento Fatorial 2^k , onde são definidos

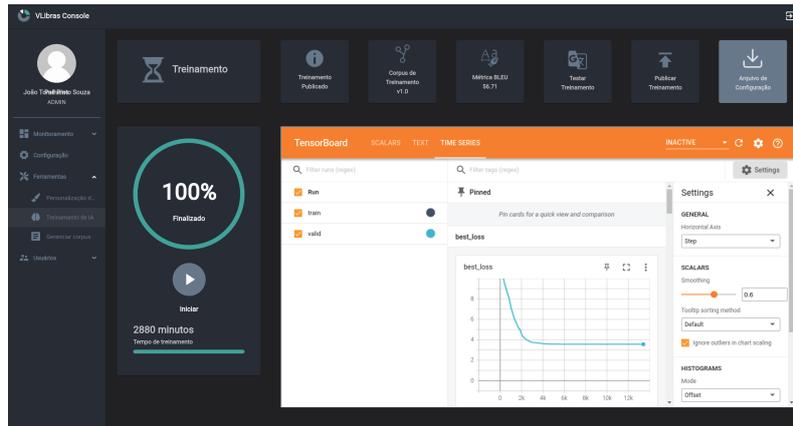


Figura 6. Interface do VLibras-Console de Visualização de Treinamento do Tradutor VLibras

k fatores de testes que podem influenciar no resultado final. Para cada fator, são definidos dois níveis, e, a partir disso, são planejados 2^k experimentos distintos variando-se esses fatores.

3.1. Fatores

Para o experimento, foram considerados os seguintes fatores que podem influenciar diretamente nos resultados: as arquiteturas de Redes Neurais utilizadas pelo Tradutor VLibras e a Quantidade de Sentenças Adicionadas pelo Automation para realizar um novo treino.

Segundo [Veríssimo et al. 2019], as abordagens *LightConv* [Wu et al. 2019] e *Transformer* [Vaswani et al. 2017] possuem variações arquiteturais que são baseadas na dimensionalidade das camadas (i.e. camadas de entrada, saída e escondidas). Tais variações podem resultar em modelos mais profundos e modelos mais simples, conhecidos como modelos *baseline*.

A quantidade de sentenças adicionadas está diretamente ligado com número de novos símbolos que farão parte do corpus. Desse modo, foram utilizados como níveis: (i) uma adição mínima, contendo apenas 100 sentenças, e (ii) uma adição maior, contendo aproximadamente 5 mil sentenças. Esses valores foram definidos empiricamente, estimando pequenas colaborações e grandes colaborações pela comunidade surda na ferramenta WikiLibras. A Tabela 2 apresenta a matriz de planejamento construída a partir dos fatores determinados.

Tabela 2. Matriz de Planejamento

Nº do Experimento	Arquitetura	Sentenças
1	<i>Transformer</i>	100
2	<i>Transformer</i>	5K
3	<i>LightConv</i>	100
4	<i>LightConv</i>	5K

3.2. Corpus bilíngue

Para validar o trabalho proposto, neste estudo de caso, foi utilizado o corpus de código aberto bilíngue Português-Libras da Suíte VLibras. Essa base conta com cerca de 67 mil sentenças em Português e a sua representação equivalente em formato glosa⁶. Essa base de dados foi dividida em três subamostras. Treinamento com aproximadamente 55 mil sentenças, Validação e Teste com 6 mil sentenças cada.

3.3. Métrica de Interesse

Os resultados foram avaliados utilizando a métrica de comparação textual BLEU [Papineni et al. 2002]. Essa métrica é bastante utilizada na literatura para avaliar modelos computacionais de tradução automática.

De acordo com *Google Cloud Translation*⁷, os valores da métrica BLEU variam de 0 a 100 e podem ser interpretados de acordo com a Tabela 3.

Tabela 3. Escala da métrica BLEU

BLEU Score	Interpretação
< 10	Sem utilidade
10 - 19	Difícil de entender a essência
20 - 29	A essência é clara, mas tem erros gramaticais significativos
30 - 40	Compreensível
40 - 50	Tradução de alta qualidade
50 - 60	Tradução de alta qualidade, fluência e adequação
> 60	Tradução geralmente melhor que a humana

A avaliação utilizando BLEU é realizada por intermédio de uma precisão de n -gramas (i.e. sequência de n letras ou palavras). Para esse trabalho, utilizou-se o BLEU 4-gramas, que é o valor de n -grama geralmente mais utilizado na literatura, como, por exemplo, os trabalhos de [Veríssimo et al. 2019] e [Lima 2015].

3.4. Configuração do Ambiente

Todos os testes foram realizados utilizando *Google Colaboratory* (Colab) [Bisong 2019]. O *Colab* é um serviço de notebook *Jupyter* hospedado que não requer configuração para ser usado, além de fornecer acesso gratuito a recursos de computação, incluindo GPUs⁸.

4. Resultados e Discussões

Nessa seção serão apresentados os resultados e uma análise dos experimentos definidos na Tabela 2. A Tabela 4 mostra os resultados dos experimentos em relação a métrica BLEU e a evolução para cada treinamento. Esta evolução é apenas a diferença entre o treinamento antes e depois da alimentação do corpus feito pelo Serviço de Integração.

Com objetivo de calcular a diferença de evolução entre os treinamentos, foram executados casos bases utilizando as duas arquiteturas. O resultado obtido foi *LightConv*

⁶A notação glosa corresponde a uma representação textual da sentença na estrutura gramatical da Libras

⁷<https://cloud.google.com/translate/automl/docs/evaluate>

⁸<https://research.google.com/colaboratory/faq.html>

com **37,75** pontos e a *Transformer* **29,84** pontos, ambos calculados de acordo com a métrica BLEU.

De acordo com a Tabela 4, é possível observar que os melhores valores de BLEU foram obtidos utilizando a arquitetura *LightConv*. Por outro lado, os experimentos com a *Transformer* apresentaram uma evolução mínima ou até mesmo negativa. O melhor caso dentre os experimentos planejados, foi com a utilização da arquitetura *LightConv* e a adição máxima de 5 mil novas sentenças pelo Serviço de Integração.

Tabela 4. Resultados da Matriz de Experimentos utilizando a métrica BLEU

Nº do Experimento	BLEU	Baseline	Evolução
1	29,06	29,84	- 0,78
2	29,99	29,84	+ 0,15
3	38,05	37,75	+ 0,30
4	39,26	37,75	+ 1,51

Por fim, a Figura 7 apresenta um gráfico evolutivo utilizando a métrica BLEU a cada iteração utilizando a Tabela 4. Considerando a primeira iteração o caso base, a segunda sendo a adição de 100 novas sentenças e na terceira 5 mil. Como é possível observar no gráfico, a arquitetura *Transformer* apresenta uma evolução inferior a *LightConv*.

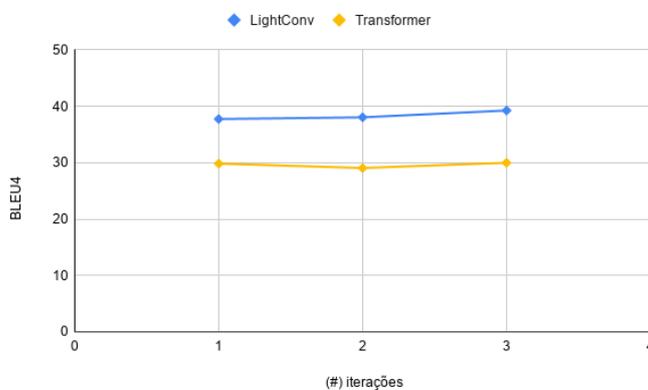


Figura 7. Gráfico Evolutivo da Métrica BLEU do Tradutor VLibras

5. Considerações Finais

Este trabalho apresentou uma ferramenta de automação (Automation) para o treinamento de forma incremental do tradutor automático VLibras. A ferramenta foi integrada com o sistema colaborativo de traduções WikiLibras para adição de novas sentenças ao corpus do tradutor de forma automática, possibilitando a cada iteração treinamentos cada vez mais robustos. Para validar a solução, utilizou-se o planejamento de experimento Fatorial 2^k definindo dois fatores que poderiam influenciar a performance do tradutor. Os resultados apontaram uma superioridade do modelo *LightConv* obtendo o resultado de **39,26** pontos na métrica BLEU com evolução **+1,5** ponto em relação ao *baseline*.

Como trabalhos futuros, pretende-se explorar o reuso dos pesos dos modelos já treinados e investigar o uso de heurísticas na sua inicialização com o objetivo de maximizar o desempenho do tradutor analisando os fatores de tempo de treinamento e qualidade da tradução.

Referências

- Araujo, T. (2012). *Uma solução para geração automática de trilhas em Língua Brasileira de Sinais em conteúdos multimídia*. PhD thesis.
- Bisong, E. (2019). *Google Colaboratory*, pages 59–64. Apress, Berkeley, CA.
- Krishna, S., Jindal, A. R., R, M., K, R., and Jayagopi, D. (2020). Virtual indian sign language interpreter. In *Proceedings of the 2020 4th International Conference on Vision, Image and Signal Processing, ICVISIP 2020*, New York, NY, USA. Association for Computing Machinery.
- Lima, M. A. C. B. (2015). *Tradução automática com adequação sintático-semântica para LIBRAS*. PhD thesis.
- Liu, D., Ma, N., Yang, F., and Yang, X. (2019). A survey of low resource neural machine translation. In *2019 4th International Conference on Mechanical, Control and Computer Engineering (ICMCCE)*, pages 39–393.
- Ott, M., Edunov, S., Baevski, A., Fan, A., Gross, S., Ng, N., Grangier, D., and Auli, M. (2019). fairseq: A fast, extensible toolkit for sequence modeling. In *Proceedings of NAACL-HLT 2019: Demonstrations*.
- Papineni, K., Roukos, S., Ward, T., and Zhu, W.-J. (2002). Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pages 311–318, Philadelphia, Pennsylvania, USA. Association for Computational Linguistics.
- Rosa Zucolotto, M. P., Ruiz, L. R., and Pinheiro, N. F. (2019). Reflexões sobre linguagem, sociedade e surdez. *Revista Uniabeu*, 12(30):134–147.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., and Polosukhin, I. (2017). Attention is all you need.
- Veríssimo, V., Silva, C., Hanael, V., Moraes, C., Costa, R., Maritan, T., Aschoff, M., and Gaudêncio, T. (2019). A study on the use of sequence-to-sequence neural networks for automatic translation of brazilian portuguese to libras. In *Proceedings of the 25th Brazillian Symposium on Multimedia and the Web, WebMedia '19*, page 101–108, New York, NY, USA. Association for Computing Machinery.
- Vichyaloetsiri, T., Wuttidittachotti, P., and (2017). Web service framework to translate text into sign language. In *2017 International Conference on Computer, Information and Telecommunication Systems (CITS)*, pages 180–184.
- Wu, F., Fan, A., Baevski, A., Dauphin, Y. N., and Auli, M. (2019). Pay less attention with lightweight and dynamic convolutions. *CoRR*, abs/1901.10430.