

AJC

Academic Journal on Computing, Engineering and Applied Mathematics

EAM

2021

Volume 2 Issue 1

Academic Journal on Computing, Engineering and Applied Mathematics



Image Designed by senivpetro / Freepik

Universidade Federal do Tocantins

Reitor

Prof. Dr. Luís Eduardo Bovolato

Vice-Reitor

Prof. Dr. Marcelo Leineker Costa

Pró-Reitoria de Graduação

Prof. Dr. Eduardo José Cezari

Pró-Reitoria de Pesquisa e Pós-Graduação

Prof. Dr. Raphael Sanzio Pimenta

Pró-Reitoria de Extensão e Cultura

Profa. Dra. Maria Santana Ferreira dos Santos

Pró-Reitoria de Administração e Finanças

Me. Carlos Alberto Moreira de Araújo Júnior

Pró-Reitoria de Assuntos Estudantis e Comunitários

Prof. Dr. Kherlley Caxias Batista Barbosa

Pró-Reitoria de Avaliação e Planejamento

Prof. Dr. Eduardo Andrea Lemus Erasmo

Pró-reitoria de Gestão e Desenvolvimento de Pessoas

Profa. Dra. Vânia Maria de Araújo Passos

Pró-Reitoria de Tecnologia da Informação e Comunicação

Prof. Dr. Ary Henrique Morais Oliveira

Direção do Campus de Palmas

Prof. Dr. Moisés de Souza Arantes Neto

Coordenação do Curso de Ciência da Computação

Prof. Dr. Eduardo Ferreira Ribeiro

Dados Internacionais de Catalogação na Publicação (CIP)

Academic Journal on Computing, Engineering and Applied Mathematics (AJCEAM) [recurso eletrônico] / Universidade Federal do Tocantins, Curso de Ciência da Computação. – vol. 02, n. 01 ([setembro/abril], 2021) – Palmas - TO, UFT, 2021. ISSN nº 2675-3588.

Quadrimestral no primeiro ano de publicação 2020

Semestral.

Disponível em:

<https://sistemas.uft.edu.br/periodicos/index.php/AJCEAM/index>

1. Ciência da Computação - periódico. 2. Matemática Aplicada. 3. Computação Aplicada. 4. Engenharias. 5. Ciências Exatas. I. Universidade Federal do Tocantins.

CDD 22.ed. 004

Ficha Catalográfica elaborada por Edson de Sousa Oliveira – CRB/2 – 1069.

Expediente

Editor-Chefe

Dr. Rafael Lima de Carvalho (UFT), Brasil

Editores

Dr. Edeilson Milhomem Silva (UFT), Brasil

Dr. Marcos Antônio Estremeto (ETEC-SP), Brasil

Dr. Rafael Lima de Carvalho (UFT), Brasil

Me. Tiago da Silva Almeida (UFT), Brasil

Dr. Warley Gramacho da Silva (UFT), Brasil

Realização

Fundação Universidade Federal do Tocantins (UFT)

Quadra 109 Norte, Avenida NS-15, ALCNO-14 | Bloco III | sala 214 |Plano Diretor Norte | 77001-090 |
Palmas / TO | Brasil

Periodicidade

Este periódico possui periodicidade semestral e utiliza a Licença Creative Commons 4.0 - CC BY-NC 4.0. Contudo, a publicação dos artigos em modalidade avançada ou ahead of print, ou seja, tão logo os manuscritos aprovados sejam editados para publicação, é possível. O AJCEAM não possui taxas de publicação, tanto pouco de submissão de manuscritos, sendo totalmente gratuita para autores e leitores.

Indexadores

Google Acadêmico, desde 9 de maio de 2020

International Standard Serial Number – ISSN, desde 28 de maio de 2020

Crossref, desde 7 de junho de 2020

Revistas de Livre Acesso – LivRe, desde 24 de junho de 2020

Sumário

1	Editorial (Português): Academic Journal on Computing, Engineering and Applied Mathematics	vi
	CARVALHO	
2	Day-Ahead Hourly Operation Planning of Distribution Networks with Photovoltaic Generation Integrated by Smart Inverters via Mixed-Integer Volt/Var Optimization Problems	1
	BASTOS ET. AL.	
3	Evaluation of a Sliding Window mechanism as Data Augmentation over Emotion Detection on Speech	11
	CARVALHO, ALMEIDA E ALMEIDA	

Editorial (Português): Academic Journal on Computing, Engineering and Applied Mathematics

Rafael Lima de Carvalho¹

¹ Universidade Federal do Tocantins, Palmas / TO, Brasil

A equipe editorial do AJCEAM vem trabalhando intensamente para melhorar a estruturação, organização e definição de objetivos para o crescimento e divulgação do periódico. Outrossim, sempre zelando pela qualidade dos artigos publicados. Por esta razão, este número possui dois importantes estudos conduzidos e revisados com excelência por nosso time qualificado de revisores.

O trabalho, intitulado “*Day-Ahead Hourly Operation Planning of Distribution Networks with Photovoltaic Generation Integrated by Smart Inverters via Mixed-Integer Volt/Var Optimization Problems*”, escrito por da Bastos et. al. [1], aborda a modelagem e a resolução de um problema de OVV (Otimização Volt/Var) que busca minimizar os desvios de magnitude de tensão em Redes de Distribuição (RD) para o planejamento de véspera da operação horária com Geração Fotovoltaica Distribuída (GFD) integrada por inversores inteligentes para um fator de potência unitário.

Além disso, Carvalho, Almeida e Almeida [2], através de seu trabalho intitulado “*Evaluation of a Sliding Window mechanism as Data Augmentation over Emotion Detection on Speech*”, avaliam a utilização de um mecanismo de janela deslizante como técnica de data augmentation no contexto de detecção de emoção em discursos. O método foi aplicado a três conjuntos de dados públicos: EMO-DB, SAVEE e RAVEDESS. Os experimentos reportam eficácia no aumento das taxas de reconhecimento de cerca de 11,95% no EMO-DB base, 22,76% no SAVEE e 18,82% no RAVEDESS quando comparado a outras abordagens da literatura.

Por fim, agradecer aos editores Tiago da Silva Almeida e Warley Gramacho da Silva pela brilhante condução dos trabalhos durante as revisões, permitindo que os esforços dos autores fossem traduzidos nestes belíssimos trabalhos. Portanto, boa leitura e excelente aprendizado.

REFERÊNCIAS

- [1] J. V. ALVES DE AZEVEDO BARROS, A. Flamarion Querubini Gonçalves, C. Francisco do Nascimento, and G. Guimarães Lage, “Day-ahead hourly operation planning of distribution networks with photovoltaic generation integrated by smart inverters via mixed-integer volt/var optimization problems.” *Academic Journal on Computing, Engineering and Applied Mathematics*, vol. 2, no. 1, p. 1–10, abr. 2021. [Online]. Available: <https://sistemas.uft.edu.br/periodicos/index.php/AJCEAM/article/view/11285>
- [2] M. A. Farias da Silva, R. L. de Carvalho, and T. d. S. Almeida, “Evaluation of a sliding window mechanism as dataaugmentation over emotion detection on speech,” *Academic Journal on Computing, Engineering and Applied Mathematics*, vol. 2, no. 1, p. 11–18, abr. 2021. [Online]. Available: <https://sistemas.uft.edu.br/periodicos/index.php/AJCEAM/article/view/11737>

Day-Ahead Hourly Operation Planning of Distribution Networks with Photovoltaic Generation Integrated by Smart Inverters via Mixed-Integer Volt/Var Optimization Problems.

João Vitor A. de A. Barros¹, Amílcar F. Q. Gonçalves¹, Claudionor F. do Nascimento¹ e
Guilherme G. Lage¹

¹ Departamento de Engenharia Elétrica, Universidade Federal de São Carlos, São Carlos, SP, 13565-905, Brasil.

Data de recebimento do manuscrito: 02/01/2021

Data de aceitação do manuscrito: 05/04/2021

Data de publicação: 14/04/2021

Abstract— The penetration of distributed photovoltaic generation in distribution networks has become a reality; however, it still imposes challenges to be overcome, such as the operation of these networks with minimal deviations in voltage magnitude profiles. Thus, this work proposes the day-ahead hourly operation planning of distribution networks aiming at minimizing the deviation in voltage magnitude profiles by means of optimal tap ratio adjustments of transformers at substations, optimal settings of remote switchable capacitor banks along primary feeders, and optimal active and reactive power injections from the distributed photovoltaic generation by smart inverters. In such Volt/Var Optimization (VVO) problem, discrete control variables associated with tap ratios of transformers and equivalent susceptances of switchable capacitor banks are handled by sinusoidal penalty functions that modify the original mixed-integer nonlinear programming problem into a nonlinear programming problem with only continuous decision variables. Results obtained for distribution networks with 69 and 135 nodes show that the proposed VVO model and its resolution methodology are effective and efficient in reducing the deviation in voltage magnitude profiles whilst also reducing the distribution network overall reactive power demand.

Keywords—mixed-integer nonlinear programming, distributed photovoltaic generation, smart inverters, Volt/Var optimization.

I. INTRODUÇÃO

O uso racional de recursos econômicos, técnicos e naturais na operação dos sistemas de energia elétrica é intrínseco às boas práticas de eficiência energética. Assim, a Geração Fotovoltaica Distribuída (GFD) integrada às Redes de Distribuição (RDs) tem se tornado uma opção viável para a redução da demanda por novos empreendimentos de grande porte devido à sua viabilidade econômica, aos avanços tecnológicos na fabricação de painéis solares fotovoltaicos cada vez mais eficientes e, não menos importante, às suas características ambientais [1, 2]. Entretanto, a presença massiva da GFD em RDs pode resultar em desvios nos perfis de magnitude de tensão, em variações nos fluxos de potência ativa e reativa nos alimentadores primários e secundários das RDs, e no aumento da sensibilidade da proteção [3]. Em [4, 5] fica evidente que a inserção de diferentes fontes de Geração Distribuída (GD), como, por exemplo, a fotovoltaica, pode reduzir custos técnico-operacionais das RDs; porém, a alo-

cação e o dimensionamento da capacidade instalada são determinantes para se atingir esse objetivo [6, 7].

Nesse contexto, tornam-se essenciais a proposição e o desenvolvimento de novos modelos e metodologias de resolução para o planejamento de véspera da operação horária das RDs com GFD. A Otimização Volt/Var (OVV) se apresenta, portanto, como a melhor ferramenta para a definição do conjunto de tomadas de decisão ótimas para o planejamento de véspera da operação de RDs segundo os fundamentos e as filosofias de *smart grids*. Nesse problema de OVV, *taps* de transformadores abaixadores nas subestações, bancos de capacitores chaveados remotamente ao longo dos alimentadores primários, e as injeções de potência ativa e reativa da GFD integrada por inversores inteligentes [8,9] são passíveis de serem controlados para se otimizar um determinado desempenho técnico-operacional da RD.

No entanto, os *taps* dos transformadores nas subestações e as susceptâncias equivalentes dos bancos de capacitores são variáveis de controle discretas [10], e, portanto, esse problema de OVV é formulado matematicamente como um problema de Programação Não Linear Inteira Mista (PNLIM) de grande porte, estático e não convexo [11]. Segundo [12], pode-se dividir as técnicas de resolução de problemas de PNLIM em quatro categorias: *Branch-and-Bound*, *Branch-and-Cut*, *Outer Approximations* e *Decomposição de Ben-*

Dados de contato: João Vitor Barros, Amílcar Gonçalves, Claudionor Nascimento and Guilherme Lage, joaobarros@estudante.ufscar.br, amilcarf@ufscar.br, claudionor@ufscar.br and glage@ufscar.br

ders generalizada. Entretanto, essas técnicas demandam alto custo computacional para problemas de grande porte em que o número de variáveis de controle discretas é elevado [11, 13, 14].

Para se resolver problemas de OVV e, assim, garantir a robustez dessa ferramenta, existem metodologias propostas na literatura correlata que não utilizam técnicas de PNLIM e propõem alternativas como técnicas de arredondamento [11], de excursão de pontos interiores à região factível delimitada pelos planos de corte [15], funções penalidade quadrática com suavização da função barreira logarítmica [16], análise de sensibilidade por variáveis duais [17], métodos de otimização global com lógica fuzzy [18], algoritmos evolutivos [19] e funções penalidade senoidal quadrática [13, 14].

Por outro lado, a integração da GFD é fundamental para operação das RDs, e seu comportamento deve ser incluído no modelo de OVV apresentado neste trabalho. Em [20] foi proposta a injeção de potência reativa por inversores inteligentes para fatores de potência não unitários, com o objetivo de minimizar o desvio nos perfis de magnitudes tensão; no entanto foram analisados resultados considerando-se a integração da GFD em um único nó. Em [21] foi apresentada a integração da GFD pela injeção de potência reativa por inversores inteligentes com fator de potência constante, também buscando minimizar os desvios de magnitude tensão. Em [22] foi apresentado um modelo para minimizar os desvios de magnitudes de tensão pela injeção de potência reativa por inversores inteligentes com fator de potência não constante, entretanto, foi utilizado uma otimização multi-objetiva com o fluxo de carga linearizado considerando apenas o ajuste nos *taps* dos transformadores com comutação sob carga.

Nesse contexto, este trabalho tem como objetivo a modelagem e a resolução de um problema de OVV que consiste na minimização dos desvios de magnitude de tensão em RDs para o planejamento de véspera da operação horária com GFD integrada por inversores inteligentes por um fator de potência unitário. Cabe ressaltar que, por se tratar de um planejamento no curto prazo, a injeção de potência na RD pela GFD pode ser considerada determinística e a modelagem dos inversores inteligentes é feita de acordo com [20]. A metodologia de tratamento das variáveis de controle discretas do problema consiste no emprego da técnica de Programação Não Linear (PNL) com funções penalidade senoidal quadrática apresentada em [13, 14].

As principais contribuições deste trabalho consistem, portanto, na proposição de um modelo atual e realista para o planejamento de véspera da operação horária das RDs com GFD integrada por inversores inteligentes e na sua resolução por uma metodologia robusta e eficiente que permite o emprego de *solvers* comerciais de problemas de PNLs de grande porte.

Um cenário foi criado para realização do estudo. As redes de distribuição são consideradas trifásicas equilibradas; a geração Fotovoltaica (FV) obedece um perfil de geração parabólico sem variações pontuais; os bancos de capacitores estão dispostos ao longo na rede com chaveamentos remoto; e a injeção de potência reativa pelo inversor inteligente pode ser tanto positiva (injetando potência reativa) quanto negativa (absorvendo potência reativa) e somente pode ocorrer no período que haja potência gerada disponível.

Este trabalho está organizado da seguinte forma. Na Seção II são apresentadas as considerações que consolida a integração dos inversores inteligentes com a rede de distribuição através da OVV. Na Seção III é apresentada a modelagem do problema de OVV para o planejamento de véspera da operação horária das RDs com GFD integrada por inversores inteligentes. Na Seção IV é apresentada a metodologia para o tratamento das variáveis de controle discretas do OVV. Na Seção V são apresentados as considerações, os resultados numéricos e as análises para RDs de 69 e 135 nós. Por fim, na Seção VI são destacadas as principais contribuições e conclusões acerca do modelo e da metodologia de resolução do problema de OVV apresentado neste trabalho.

II. INVERSORES INTELIGENTES

O conceito de redes inteligentes trabalha com a possibilidade de comunicação entre os dispositivos presentes no sistema elétrico. Há diversos trabalhos na literatura que propõem arquiteturas de comunicação para inversores inteligentes, sendo alguns voltados ao controle de magnitudes de tensão, como em [23], no qual uma infraestrutura de comunicação para o controle da magnitude da tensão em sistemas de transmissão de potência é proposta; em [24] é proposta uma arquitetura de controle de tensão descentralizada e não hierárquica baseada em agentes *fuzzy* cooperativos; e em [25] um estudo foi realizado para avaliar a performance da comunicação no controle de inversores inteligentes com o objetivo de controlar as magnitudes de tensão em redes de distribuição.

Um inversor inteligente para sistemas fotovoltaicos é um conversor de Corrente Contínua (CC) para Corrente Alternada (CA), conectado na rede de distribuição, dotado de um sistema de comunicação e que possa operar em diferentes modos de injeção de potência ativa e/ou reativa, conforme seja estabelecido pelo centro de operações da rede. Para isso, é essencial que a GFD tenha a capacidade de injetar energia na rede por meio dos painéis fotovoltaicos, bem como armazenar energia em baterias ou supercapacitores, com a finalidade de gerenciar as injeções de potência ativa e reativa nos quatro quadrantes (vide Figura 1 da referência [20]). Dessa forma, seguindo os padrões da norma IEEE 1547-2018, o controle empregado aqui na GFD permite a operação no modo OVV por meio do ajuste da magnitude e fase de tensão nos inversores inteligentes, controlando, portanto, as potências ativa e reativa, injetada ou absorvida da rede.

O sincronismo de fase, para que ocorra a integração dos inversores inteligentes à rede, é realizado por meio do algoritmo de bloqueio de fase (*Phase Locked Loop - PLL*) que produz uma referência sincronizada em fase e frequência com o sinal de entrada (tensão da rede). O PLL tem papel fundamental no rastreamento do ângulo de tensão do sistema CA no ponto de acoplamento comum.

Para controlar as potências ativa e reativa que serão fornecidas pelo inversor inteligente é utilizada a transformação de Park ($dq0$). Na transformada de Park os sinais no referencial trifásico (abc) são convertidos para um referencial de duas coordenadas síncronas ($dq0$). O controle em coordenadas $dq0$ permite que os sinais passem de um sistema de característica de sinal alternado para um sistema com car-

acterística de sinal contínuo, separando o controle de potência ativa no eixo direto (d) e o controle de potência reativa no eixo de quadratura (q) [26]. Além disso, o fato dos sinais terem a característica CC permite que o controle em $dq0$ utilize controladores proporcional-integral (PI), zerando o erro de regime permanente. O leitor interessado em se aprofundar na modelagem do inversor inteligente veja a referência [26].

III. MODELAGEM DO PROBLEMA DE OVV

Dado que problemas de OVV são formulados como problemas de PNLIM de grande porte, estáticos e não convexos [11], tais problemas podem ser expressos matematicamente pela seguinte formulação genérica:

$$\min Z = \sum_{t=1}^T f(x^{(t)}, y^{(t)}) \quad (1a)$$

$$\text{s.a: } g(x^{(t)}, y^{(t)}) = 0, \quad \forall t \in \mathcal{T} \quad (1b)$$

$$h(x^{(t)}, y^{(t)}) \leq 0, \quad \forall t \in \mathcal{T} \quad (1c)$$

$$x_{\min} \leq x^{(t)} \leq x_{\max}, \quad \forall t \in \mathcal{T} \quad (1d)$$

$$y^{(t)} \in \mathcal{D}_y, \quad \forall t \in \mathcal{T} \quad (1e)$$

em que t representa cada intervalo de discretização do horizonte de planejamento $\mathcal{T} = \{1, \dots, T\}$; $x^{(t)} \in \mathbb{R}^m$ representa o vetor das variáveis de decisão contínuas do problema, cujos limites inferiores e superiores são dados, respectivamente, pelos vetores $x_{\min} \in \mathbb{R}^m$ e $x_{\max} \in \mathbb{R}^m$, $y^{(t)} \in \mathbb{R}^n$ representa o vetor das variáveis de decisão discretas do problema, cujos valores são definidos pelo conjunto \mathcal{D}_y ; A função objetivo $Z: \mathbb{R}^m \cup \mathbb{R}^n \mapsto \mathbb{R}$ consiste no somatório dos desvios quadráticos das magnitudes de tensão nodais da RD; $g: \mathbb{R}^m \cup \mathbb{R}^n \mapsto \mathbb{R}^r$ representa as restrições de igualdade do problema e consiste nas equações de balanço de potência ativa e reativa nos nós da RD, com $r < m + n$; e, por fim, $h: \mathbb{R}^m \cup \mathbb{R}^n \mapsto \mathbb{R}^s$ representa as restrições de desigualdade funcionais do problema e consiste na modelagem das injeções de potência ativa e reativa dos inversores inteligentes que integram a GFD à RD.

A modelagem do problema de planejamento de véspera da operação horária de RDs com GFD integrada por inversores inteligentes é detalhada nas próximas subseções.

a. Função objetivo

A função objetivo considerada neste trabalho consiste no somatório dos desvios quadráticos das magnitudes de tensão nodais da RD para todos os intervalos de discretização do horizonte de planejamento, e pode ser matematicamente expressa por:

$$Z = \sum_{t=1}^T \sum_{k \in \mathcal{C}} (1 - V_k^{(t)})^2 \quad (2)$$

em que $V_k^{(t)}$ é a magnitude de tensão no nó k no intervalo de discretização t do horizonte de planejamento \mathcal{T} ; e \mathcal{C} é o conjunto dos nós de carga da RD.

b. Restrições de igualdade

As restrições de igualdade consistem nas equações de balanço de potência ativa e reativa nos nós da RD, e podem ser matematicamente expressas por:

$$P_k^{(t)} + P_k^{fv(t)} - \sum_{m \in \Omega_k} P_{km}^{(t)}(V^{(t)}, \theta^{(t)}, t_{km}^{(t)}) = 0, \quad \forall k \in \mathcal{C}, \forall t \in \mathcal{T} \quad (3a)$$

$$Q_k^{(t)} + Q_k^{fv(t)} + b_k^{sh(t)} V_k^{(t)2} - \sum_{m \in \Omega_k} Q_{km}^{(t)}(V^{(t)}, \theta^{(t)}, t_{km}^{(t)}) = 0, \quad \forall k \in \mathcal{C}, \forall t \in \mathcal{T} \quad (3b)$$

$$S_k^{fv(t)} = \sqrt{P_k^{fv(t)2} + Q_k^{fv(t)2}}, \quad \forall k \in \mathcal{F}, \forall t \in \mathcal{T} \quad (3c)$$

em que $V^{(t)}$ e $\theta^{(t)}$ são variáveis contínuas associadas às magnitudes e aos ângulos de fase das tensões nodais; $P^{fv(t)}$ e $Q^{fv(t)}$ são variáveis contínuas associadas às injeções de potência ativa e reativa por inversores inteligentes que integram a GFD à RD (os nós que não possuem GFD integrada por inversores inteligentes têm $P^{fv(t)}$ e $Q^{fv(t)}$ iguais a zero); $t^{(t)}$ e $b^{sh(t)}$ são variáveis de controle discretas associadas, respectivamente, aos *taps* de transformadores abaixadores nas subestações e às susceptâncias equivalentes dos bancos de capacitores chaveados remotamente ao longo dos alimentadores primários; $P^{(t)}$ e $Q^{(t)}$ são as injeções líquidas de potência ativa e reativa nos nós da RD; $S^{fv(t)}$ é um parâmetro determinístico associado à injeção de potência na RD pela GFD não superior à potência nominal do inversor inteligente; e $P_{km}^{(t)}$ e $Q_{km}^{(t)}$ são os fluxos de potência ativa e reativa nas linhas entre os nós k e m .

As restrições (3a) e (3b) são as equações tradicionais que modelam os fluxos de potência ativa e reativa nos nós da RD. As restrições (3c) modela a injeção de potência ativa e reativa na RD pela integração da GFD por inversores inteligentes. Por fim, \mathcal{F} é o conjunto dos nós da RD com GFD integrada por inversores inteligentes; Ω_k é o conjunto de todos os nós conectados ao nó k .

c. Limites operacionais e valores discretos de variáveis de controle

As restrições de desigualdade consistem, respectivamente, na restrição de funcional de desigualdade (3c) que modela a injeção de potência ativa e reativa da GFD por inversores inteligentes de acordo com [20], e as restrições (4c), (4d) e (4e) são associadas aos limites técnico-operacionais da RD:

$$0 \leq P_f^{(t)} \leq S_f^{(t)}, \quad \forall f \in \mathcal{F}, \forall t \in \mathcal{T} \quad (4a)$$

$$|Q_f^{(t)}| \leq S_f^{(t)}, \quad \forall f \in \mathcal{F}, \forall t \in \mathcal{T} \quad (4b)$$

$$V_{k,\min} \leq V_k^{(t)} \leq V_{k,\max}, \quad \forall k \in \mathcal{C}, \forall t \in \mathcal{T} \quad (4c)$$

$$t_{km}^{(t)} \in \mathcal{D}_t, \quad \forall \{k, m\} \in \mathcal{J}, \forall t \in \mathcal{T} \quad (4d)$$

$$b_k^{sh(t)} \in \mathcal{D}_{b^{sh}}, \quad \forall k \in \mathcal{B}^{sh}, \forall t \in \mathcal{T} \quad (4e)$$

em que \mathcal{J} é o conjunto dos ramos associados a transformadores abaixadores nas subestações; e \mathcal{B}^{sh} é o conjunto

dos nós com bancos de capacitores chaveados remotamente ao longo dos alimentadores primários.

IV. TRATAMENTO DAS VARIÁVEIS DE CONTROLE DISCRETAS

Em [13, 14], uma função penalidade senoidal quadrática foi incorporada à função objetivo de um problema de despacho ótimo de reativos com objetivo de impedir que as variáveis discretas assumissem valores não discretos. Essa função penalidade senoidal quadrática $\phi : \mathbb{R} \mapsto [0, 1]$ assume valor zero quando $y_k^{(t)}$ assume um valor discreto e pode ser expressa matematicamente por:

$$\phi(y_k^{(t)}) = \text{sen}^2 \left[\left(\frac{y_k^{(t)}}{s_k^{(t)u} - s_k^{(t)l}} \right) \pi + \alpha_k \right] \quad (5)$$

em que $s_k^{(t)u}$ é o valor discreto imediatamente superior a $y_k^{(t)}$; $s_k^{(t)l}$ é o valor discreto imediatamente inferior a $y_k^{(t)}$; e α_k é uma constante definida no intervalo $[0, \pi]$ tal que a função penalidade senoidal quadrática se anule para valores de $y^{(t)}$ definidos em \mathcal{D}_y . A constante α_k é definida por:

$$\alpha_k = \left(\left\lceil \frac{s_k^{(t)u}}{s_k^{(t)u} - s_k^{(t)l}} \right\rceil - \left(\frac{s_k^{(t)l}}{s_k^{(t)u} - s_k^{(t)l}} \right) \right) \pi \quad (6)$$

em que $\lceil \cdot \rceil$ é a função *ceil* que arredonda qualquer número real para o inteiro superior imediato.

Dessa forma, problemas de OVV formulados como problemas de PNLIM podem ser transformados em problemas de PNL equivalentes da seguinte forma:

$$\min Z + \gamma \sum_{t=1}^T \sum_{k=1}^n \phi(y_k^{(t)}) \quad (7a)$$

$$\text{s.a: } g(x^{(t)}, y^{(t)}) = 0, \quad \forall t \in \mathcal{T} \quad (7b)$$

$$h(x^{(t)}, y^{(t)}) \leq 0, \quad \forall t \in \mathcal{T} \quad (7c)$$

$$x_{\min} \leq x^{(t)} \leq x_{\max}, \quad \forall t \in \mathcal{T} \quad (7d)$$

$$y_{\min} \leq y^{(t)} \leq y_{\max}, \quad \forall t \in \mathcal{T} \quad (7e)$$

em que γ é um parâmetro maior que zero associado à amplitude da função penalidade senoidal quadrática; y_{\min} é igual a $\min(\mathcal{D}_y)$ e y_{\max} é igual a $\max(\mathcal{D}_y)$.

O algoritmo para a resolução de problemas de PNLIM transformados em problemas de PNL equivalentes apresentado neste trabalho foi adaptado de [13, 14]. A cada iteração, o valor γ é incrementado pelo fator c , fixo no intervalo $(1, 10]$, para que a diferença entre o valor calculado e o discreto seja menor ou igual a uma tolerância ξ . O algoritmo utilizado neste trabalho é apresentado em Algoritmo 1.

V. RESULTADOS NUMÉRICOS E DISCUSSÕES

Os testes aqui apresentados foram realizados para a RD de 69 nós adaptada de [27], com tensão base $V_{base} = 12,66$ kV, e a RD de 135 nós adaptada de [28], com tensão

Algorithm 1 Algoritmo de Tratamento das Variáveis de Controle Discretas.

```

1: Início
2: Defina  $\gamma^0$ ,  $\xi$  e  $c$ .
3: Faça  $k \leftarrow 0$ .
4: Faça  $flag^k \leftarrow \text{FALSO}$  (parâmetro booleano de controle).
5: enquanto  $flag^k = \text{FALSO}$  faça
6:   Faça  $k \leftarrow k + 1$ .
7:   Resolva (7).
8:   se  $|y^{(t)} - y_{disc}^{(t)}| \leq \xi$  então
9:     Faça  $flag^k \leftarrow \text{VERDADEIRO}$ .
10:    retorna Solução encontrada:  $(x^{(t)}, y_{disc}^{(t)})$ .
11:  senão
12:    Faça  $\gamma^{k+1} \leftarrow c\gamma^k$ .
13:  fim se
14: fim enquanto
15: end

```

base de $V_{base} = 13,80$ kV, ambas com potência base $S_{base} = 1000$ kVA. Os perfis de carga das classes de consumidores ao longo do horizonte de planejamento considerados para os nós das RDs foram baseados em [10], de acordo com a Figura 1 e a Tabela 1. Um ramo 0 – 1 foi adicionado a cada RD para modelar a atuação de transformadores abaixadores na subestação, os quais possuem *taps* com limites de 0,90 e 1,10 por unidade (p.u.), com passos de discretização de 0,00625 p.u. Os limites das magnitudes de tensão são 0,95 e 1,05 p.u., e a tensão no nó 0 (enrolamento primário do transformador abaixador na subestação) é fixa e igual a 1,00 p.u. com fase 0 rad.

Os testes foram realizados em um notebook ACER E5-574-78LR, com drive SSD de 512 GB, processador Intel Core i7-6500 com 3.1 GHz e Memória Ram de 8 GB.

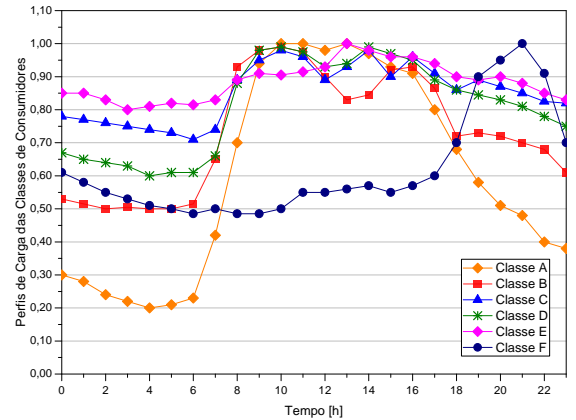


Figura 1: Perfis de carga das classes de consumidores ao longo do horizonte de planejamento [10].

De acordo com [29], o estudo de impactos da GFD nas RDs pode considerar uma penetração fotovoltaica equivalente a aproximadamente 30% da carga da RD. Dessa forma, os nós com conexão de GFD na RD de 69 nós são: 7, 16, 21, 26, 33, 39, 48, 49, 51, 53, e 62, cujos inversores inteligentes possuem uma capacidade nominal de 0,1 p.u.; e os nós com conexão de GFD na RD de 135 nós são: 7, 18, 24, 28, 30, 37, 42, 53, 58, 67, 74, 85, 90, 95, 106, 110, 119, 126, 128, e 134, cujos inversores inteligentes possuem uma capacidade nominal de 0,5 p.u.. Os perfis de injeção de potência por cada conjunto de geração fotovoltaica para ambas as RDs são apresentados nas Figuras 2 e 3.

TABELA 1: DISTRIBUIÇÃO DAS CLASSES DE CONSUMIDORES NAS RDS [10]

Classe	RD de 69 nós	RD de 135 nós
A	59, 60, 62, 63, 64, 66, 68, 69.	3, 5, 6, 9, 15, 37, 47, 61, 64, 65, 76, 80, 86, 92, 96, 102, 108, 113, 120, 125, 127, 132, 134.
B	42, 43, 44, 48, 50, 51, 53, 54, 55, 56.	4, 7, 17, 24, 34, 40, 44, 46, 53, 56, 72, 73, 77, 82, 90, 95, 98, 99, 101, 114, 118, 121, 126.
C	37, 38, 39.	2, 16, 21, 23, 30, 32, 39, 42, 52, 58, 60, 67, 83, 104, 106, 110, 119, 131.
D	40, 41, 57, 58.	49, 54, 62, 63, 68, 70, 74, 81, 85, 88, 105, 116, 123, 124.
E	28, 29, 33, 34, 35.	8, 14, 18, 38, 41, 48, 59, 71, 78, 94, 117, 128, 133, 135.
F	6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20, 21, 22, 24, 26, 27.	10, 20, 26, 28, 31, 35, 51, 66, 75, 89, 97, 100, 112, 122.

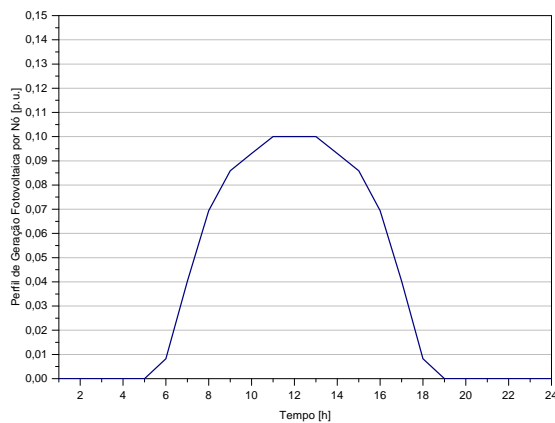


Figura 2: Perfil de injeção de potência $S^{fv(t)}$ por cada conjunto de geração fotovoltaica para a RD de 69 nós.

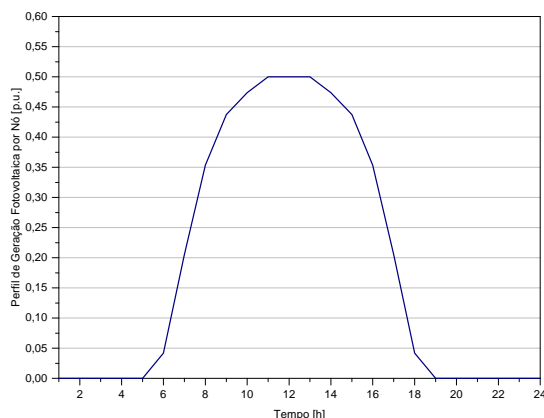


Figura 3: Perfil de injeção de potência $S^{fv(t)}$ por cada conjunto de geração fotovoltaica para a RD de 135 nós.

Por fim, para ambas RDs, os bancos de capacitores chaveados remotamente foram alocados e dimensionados de acordo com [10], os quais são formados, por sua vez, por conjuntos de 0,3 p.u. cada.

Foram realizadas três estudos de caso para a análise dos impactos da GFD integrada às RDs por inversores inteligentes: no estudo de caso 1 considerou-se a RD sem GFD (ou seja, desconsiderou-se a restrição (3c) do problema de OVV, com $P^{fv(t)}$ e $Q^{fv(t)}$ fixos e iguais a 0 p.u. para todos os

intervalos de discretização t do horizonte de planejamento); no estudo de caso 2 considerou-se apenas a injeção de potência ativa pela GFD (ou seja, desconsiderou-se a restrição (3c) do problema de OVV, com $P^{fv(t)} = S^{fv(t)}$ e $Q^{fv(t)}$ fixo e igual a 0 p.u. para todos os intervalos de discretização t do horizonte de planejamento); e no estudo de caso 3 considerou-se o modelo completo proposto neste trabalho para o problema de OVV, no qual a integração da GFD é feita por inversores inteligentes que injetam potência ativa e reativa na RD.

Esses problemas de OVV e o algoritmo de tratamento das variáveis de controle discretas apresentado no Algoritmo 1 foram implementados no ambiente de modelagem de problemas de programação matemática AMPL [30, 31] e resolvidos pelo *solver* CONOPT [32], versão 3.17A, que utiliza o método gradiente reduzido generalizado para resolver problemas de PNL.

Os parâmetros do Algoritmo 1, γ e c , determinam a amplitude da função senoidal durante as iterações do algoritmo, se o valor de c for muito alto, necessita de um menor número de iterações, entretanto, pode impedir um melhor ponto de operação da RD; se o valor de c for muito baixo aumenta o número de iterações necessárias e o custo computacional mas pode proporcionar um ponto de operação melhor. Os valores de $\gamma^0 = 0,00001$, $c = 2,8$ foi adotado empiricamente. O erro das variáveis $\xi = 0,0002$ é determinado em função do passo de ajuste das variáveis de controle.

a. Resultados Numéricos para a RD de 69 Nós

Embora a RD de 69 nós seja de pequeno porte quando comparada às RDs reais, os resultados numéricos apresentados a seguir nesta subseção serviram de base para a validação do modelo de problema de OVV proposto neste trabalho, bem como do algoritmo de tratamento das variáveis de controle discretas. O tempo de processamento do caso 1 é de 6,65625 segundos, do caso 2 é de 7,29688 segundos e do caso 3 é de 145,01628 segundos. A Tabela 2 apresenta a comparação de indicadores do desempenho operacional da RD de 69 nós obtidos para os três estudos de caso considerados neste trabalho.

TABELA 2: INDICADORES OPERACIONAIS PARA A COMPARAÇÃO DOS RESULTADOS ENTRE OS TRÊS ESTUDOS DE CASO SIMULADOS PARA A RD DE 69 NÓS

Indicador	Caso 1	Caso 2	Caso 3
$V_{med.}$ [p.u.]	0,9998	1,0002	1,0000
$V_{min.}$ [p.u.]	0,9643	0,9717	0,9728
$V_{max.}$ [p.u.]	1,0105	1,0115	1,0115
$P_{med.}^{loss}$ [kW]	205,988	175,158	203,27
$P_{0,med.}$ [kW]	2.992,41	2.961,58	2.989,67
$Q_{0,med.}$ [kVAR]	-3.217,07	-3.326,72	-3.533,76

A partir desses indicadores, pode-se perceber que o estudo de caso 3, no qual a GFD integrada à RD por inversores inteligentes é modelada e incorporada ao problema de OVV, resulta em valores médios para as magnitudes de tensão iguais a 1 p.u., com uma menor diferença entre as magnitudes de tensão máxima e mínima, evidenciando a eficácia do modelo em minimizar o desvio de magnitudes de tensão. Além disso, apesar de as perdas ativas médias na RD se

aproximarem das do estudo de caso 1, no qual a GFD não é considerada, a demanda de potência reativa da RD é menor, o que proporciona maiores margens de carregamento para o sistema elétrico de potência como um todo. Nesses estudos de caso, em específico, a RD de 69 nós está “injetando” potência reativa no sistema de transmissão.

Os perfis das magnitudes de tensão nodal são apresentados para os três estudos de caso nas Figuras 4, 5 e 6.

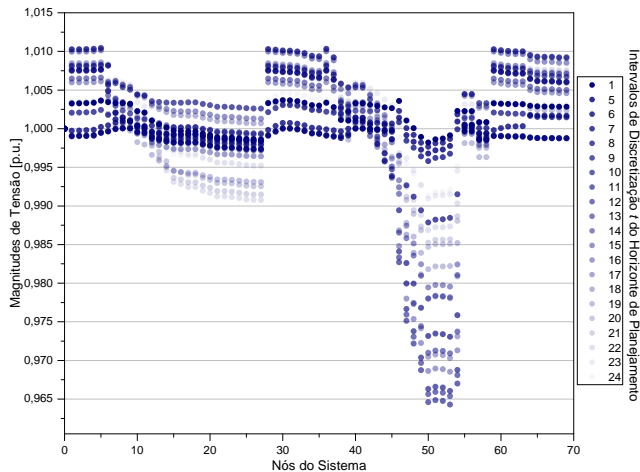


Figura 4: Magnitudes de tensão da RD de 69 nós para o estudo de caso 1.

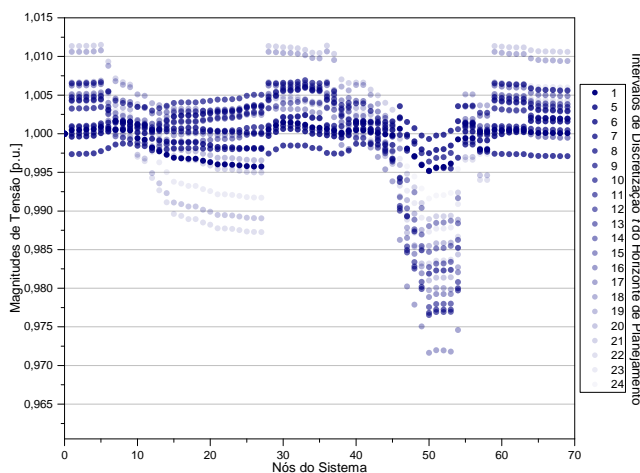


Figura 5: Magnitudes de tensão da RD de 69 nós para o estudo de caso 2.

A RD de 69 nós possui a característica de menores valores para as magnitudes de tensão nos nós 48 a 54 devidos aos respectivos perfis de carga das cargas de consumidores e, principalmente, à topologia da RD. Comparando os valores das magnitudes de tensão nesses nós, observam-se valores de magnitude de tensão com menores desvios devido à penetração da GFD na RD.

Nos três estudos de caso o algoritmo de tratamento das variáveis de controle discretas convergiu em 12 iterações; o processo de convergência de algumas dessas variáveis para o estudo de caso 3 é apresentado na Tabela 3.

O algoritmo toma como ponto inicial o estado do sistema (magnitude de tensão na barra $V\theta$, posição do *tap* do transformador na subestação e susceptâncias equivalentes

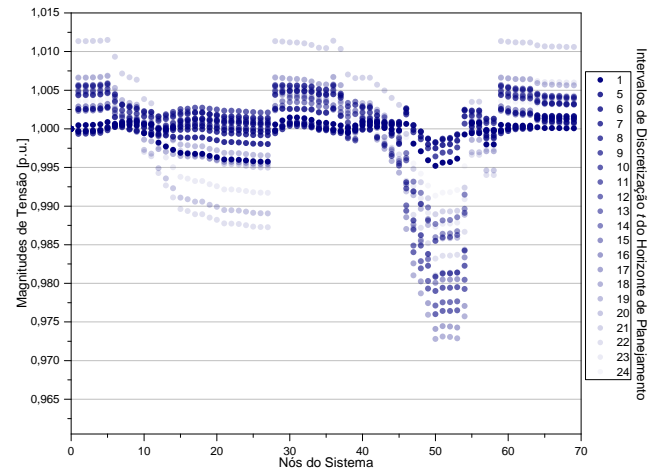


Figura 6: Magnitudes de tensão da RD de 69 nós para o estudo de caso 3.

dos bancos de capacitores chaveados remotamente), baseado nisso escolhe-se empiricamente os valores de γ^0 e c , dado que a função senoidal insere não convexidades na função objetivo. Dessa forma, a escolha desses parâmetros buscou a convergência para os menores valores da função objetivo, como apresentado na Tabela 2. Cabe ressaltar que, o ajuste ótimo das susceptâncias equivalentes *shunt* dos bancos de capacitores chaveados remotamente é sensível ao estado inicial do sistema.

A Figura 7 apresenta os valores absolutos das funções objetivo da RD de 69 nós dos 3 casos estudados comparando o problema com relaxação contínua das variáveis de controle e o problema discretizado. A maior diferença entre os casos avaliados é o período das 20 horas do caso 3 que é de 0,00079, ainda assim, as curvas de cada caso estão próximas evidenciando que a função penalidade senoidal é eficaz em discretizar as variáveis de controle.

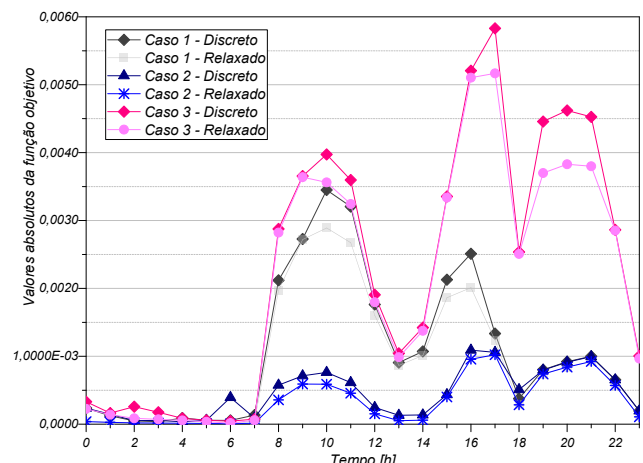


Figura 7: Valores da função objetivo dos casos avaliados do problema discreto e do problema com relaxação contínua das variáveis de controle da RD de 69 nós.

b. Resultados Numéricos para a RD de 135 Nós

Os resultados numéricos apresentados a seguir nesta subseção serviram de base para testes de robustez do modelo de problema de OVV proposto neste trabalho, devido ao seu

TABELA 3: CONVERGÊNCIA DO ALGORITMO DE TRATAMENTO DAS VARIÁVEIS DE CONTROLE DISCRETAS PARA A RD DE 69 NÓS NO ESTUDO DE CASO 3

Iteração	γ	$t_{0-1}^{(6)}$	$t_{0-1}^{(9)}$	$t_{0-1}^{(17)}$	$t_{0-1}^{(18)}$	$t_{0-1}^{(21)}$	$b_8^{sh(1)}$	$b_8^{sh(3)}$	$b_{36}^{sh(7)}$	$b_{56}^{sh(6)}$	$b_{64}^{sh(20)}$
0	0,0000100	1,05000	1,03068	1,01759	1,01588	1,01558	0,600000	0,600000	0,897987	0,600000	0,000000
1	0,0000280	1,05000	1,03112	1,01777	1,01589	1,01557	0,600000	0,596242	0,897570	0,600000	0,000000
2	0,0000784	1,05000	1,03119	1,01791	1,01595	1,01556	0,600000	0,594056	0,897719	0,599690	0,000000
3	0,0002195	1,05000	1,03125	1,01818	1,01642	1,01542	0,600000	0,592885	0,898066	0,599260	0,000000
4	0,0006147	1,05000	1,03128	1,01842	1,01789	1,01349	0,600000	0,593614	0,898830	0,598648	0,015528
5	0,0017210	1,05000	1,03126	1,01859	1,01834	1,01294	0,600000	0,596345	0,899537	0,599433	0,009327
6	0,0048189	1,05000	1,03126	1,01869	1,01859	1,01267	0,600000	0,596727	0,899579	0,599480	0,008231
7	0,0134929	1,05000	1,03125	1,01873	1,01869	1,01257	0,600000	0,596922	0,899603	0,599510	0,007727
8	0,0377802	1,05000	1,03125	1,01874	1,01873	1,01253	0,600000	0,599735	0,899965	0,599954	0,000560
9	0,1057850	1,05000	1,03125	1,01875	1,01874	1,01251	0,600000	0,599827	0,899980	0,599974	0,000327
10	0,2961970	1,05000	1,03125	1,01875	1,01875	1,01250	0,600000	0,599896	0,899988	0,599985	0,000195
11	0,8293510	1,05000	1,03125	1,01875	1,01875	1,01251	0,600000	0,599978	0,899997	0,599997	0,000043
12	2,3221800	1,05000	1,03125	1,01875	1,01875	1,01250	0,600000	0,600000	0,900000	0,600000	0,000000

maior porte e, conseqüentemente, maior número de variáveis de controle discretas. O tempo de processamento do caso 1 é de 25,8438 segundos, do caso 2 é de 19,9688 segundos e do caso 3 é de 476,047 segundos. A Tabela 4 apresenta a comparação de indicadores do desempenho operacional da RD de 135 nós obtidos para os três estudos de caso considerados neste trabalho.

TABELA 4: INDICADORES OPERACIONAIS PARA A COMPARAÇÃO DOS RESULTADOS ENTRE OS TRÊS ESTUDOS DE CASO SIMULADOS PARA A RD DE 135 NÓS

Indicador	Caso 1	Caso 2	Caso 3
$V_{med.}$ [p.u.]	0,9999	1,0010	0,9999
$V_{min.}$ [p.u.]	0,9833	0,9835	0,9818
$V_{max.}$ [p.u.]	1,0164	1,0171	1,0142
$P_{med.}^{loss}$ [kW]	378,99	226,27	326,24
$P_{0,med.}$ [kW]	13.976,11	13.823,40	13.923,36
$Q_{0,med.}$ [kVAr]	4.949,90	4.099,28	1.962,52

A partir desses indicadores, pode-se perceber que o estudo de caso 3 resulta em valores médios para as magnitudes de tensão mais próximos de 1 p.u., com uma menor diferença entre as magnitudes de tensão máxima e mínima. Além disso, apesar de as perdas ativas médias na RD se aproximarem mais das do estudo de caso 1 do que das do estudo de caso 2, a demanda de potência reativa da RD no estudo de caso 3 é significativamente menor devido à injeção de potência reativa pelos inversores inteligentes, o que proporciona maiores margens de carregamento para o sistema elétrico de potência como um todo.

Os perfis das magnitudes de tensão nodal são apresentados para os três estudos de caso nas Figuras 8, 9 e 10.

As injeções de potência ativa e reativa por inversores inteligentes ao longo do horizonte de planejamento da RD para o estudo de caso 3 são apresentadas nas Figuras 12 e 13.

A convergência do algoritmo de tratamento das variáveis de controle discretas se deu em 15 iterações para o estudo de caso 3; o processo de convergência de algumas dessas variáveis para o estudo de caso 3 é apresentado na Tabela 5.

A Figura 11 apresenta os valores absolutos das funções objetivo da RD de 135 nós dos 3 casos estudados analisando o algoritmo de discretização da penalidade senoidal. A maior diferença entre os casos avaliados é o período das 17 horas do caso 1 que é de 0,0022. O algoritmo de discretização apresenta resposta semelhante nas RDs de 69 e 135

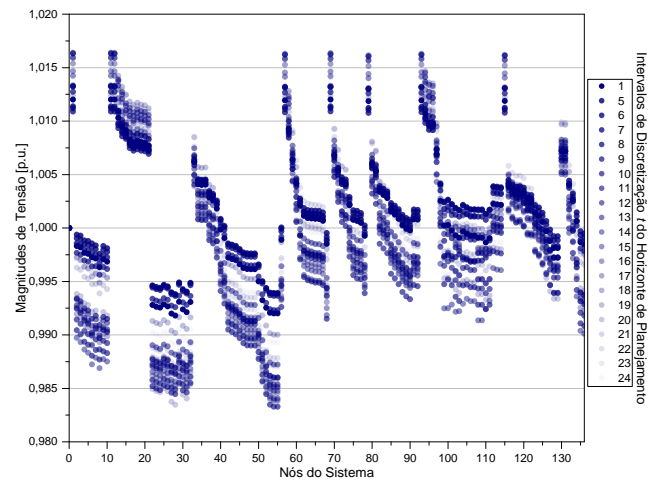


Figura 8: Magnitudes de tensão da RD de 135 nós para o estudo de caso 1.

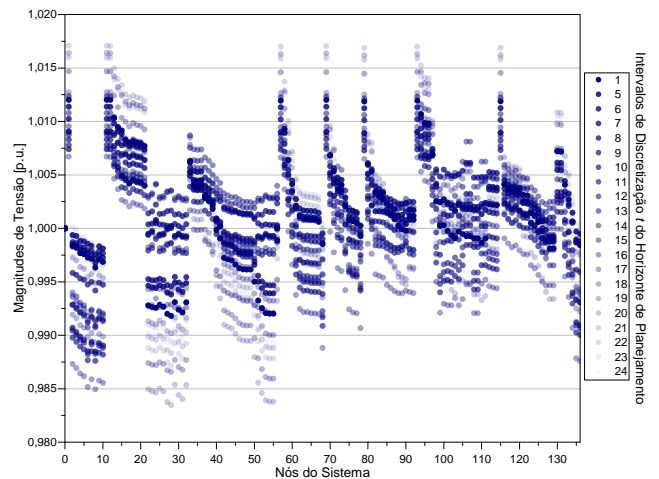


Figura 9: Magnitudes de tensão da RD de 135 nós para o estudo de caso 2.

nós, mostrando a robustez da função penalidade senoidal na metodologia empregada.

Ressalta-se, por fim, que esta proposta consiste em um avanço das propostas apresentadas em [10, 11] dada a modelagem da injeção de potência ativa e reativa por inversores inteligentes que integram a GFD. A integração da GFD por inversores inteligentes viabiliza, pois, o aumento das opções de controle das RDs. Portanto, quaisquer comparações dos

TABELA 5: CONVERGÊNCIA DO ALGORITMO DE TRATAMENTO DAS VARIÁVEIS DE CONTROLE DISCRETAS PARA A RD DE 135 NÓS NO ESTUDO DE CASO 3

Iteração	γ	$t_{0-1}^{(3)}$	$t_{0-1}^{(10)}$	$t_{0-1}^{(18)}$	$t_{0-1}^{(20)}$	$t_{0-1}^{(22)}$	$b_{41}^{sh(1)}$	$b_{69}^{sh(3)}$	$b_{74}^{sh(16)}$	$b_{111}^{sh(20)}$	$b_{127}^{sh(22)}$
0	0,0000100	0,958666	1,00554	0,942023	0,936234	0,940971	0,600000	0,286222	0,600000	0,352742	0,000000
1	0,0000280	0,958676	1,00591	0,942575	0,936407	0,941149	0,600000	0,287594	0,600000	0,332966	0,000000
2	0,0000784	0,958596	1,00613	0,942905	0,936506	0,941314	0,600000	0,287366	0,600000	0,328294	0,000515
3	0,0002195	0,958275	1,00621	0,943167	0,936617	0,941885	0,600000	0,284356	0,600000	0,323222	0,013738
4	0,0006147	0,957600	1,00623	0,943374	0,936754	0,942492	0,600000	0,281923	0,599989	0,318521	0,018320
5	0,0017210	0,957018	1,00625	0,943573	0,936976	0,942964	0,600000	0,287557	0,600000	0,313457	0,013380
6	0,0048189	0,956626	1,00625	0,943666	0,937212	0,943345	0,600000	0,293444	0,600000	0,307537	0,007446
7	0,0134929	0,956408	1,00625	0,943715	0,937372	0,943576	0,600000	0,297132	0,600000	0,303350	0,003316
8	0,0377802	0,956311	1,00625	0,943741	0,937450	0,943682	0,599999	0,298910	0,600000	0,301331	0,001300
9	0,1057850	0,956272	1,00625	0,943747	0,937481	0,943725	0,599984	0,299594	0,600000	0,300499	0,000483
10	0,2961970	0,956258	1,00625	0,943749	0,937493	0,943740	0,599993	0,299856	0,600000	0,300178	0,000172
11	0,8293510	0,956253	1,00625	0,943750	0,937498	0,943747	0,599997	0,299949	0,600000	0,300064	0,000062
12	2,3221800	0,956251	1,00625	0,943750	0,937499	0,943749	0,599999	0,299982	0,600000	0,300023	0,000022
13	6,5021100	0,956250	1,00625	0,943750	0,937500	0,943750	0,600000	0,299993	0,600000	0,300008	0,000008
14	18,205900	0,956250	1,00625	0,943750	0,937500	0,943750	0,600000	0,299998	0,600000	0,300003	0,000003
15	50,976600	0,956250	1,00625	0,993750	0,937500	0,943750	0,600000	0,300000	0,600000	0,300000	0,000000

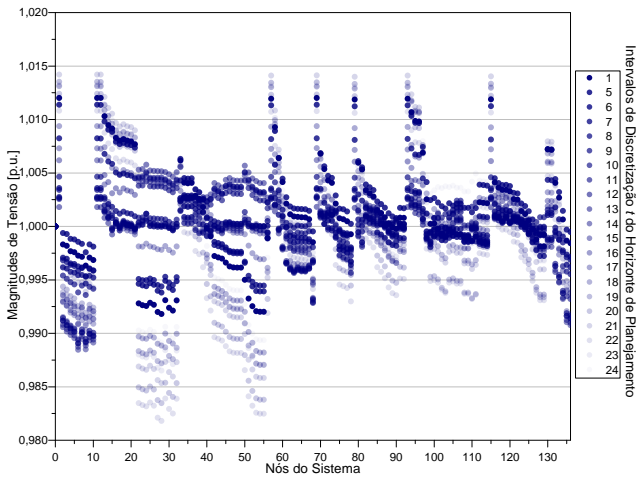


Figura 10: Magnitudes de tensão da RD de 135 nós para o estudo de caso 3.

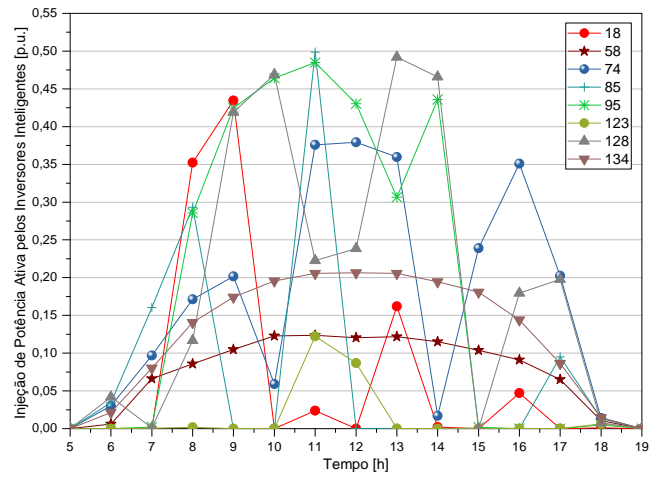


Figura 12: Perfis de injeção de potência ativa pelos inversores inteligentes da RD de 135 nós para o estudo de caso 3.

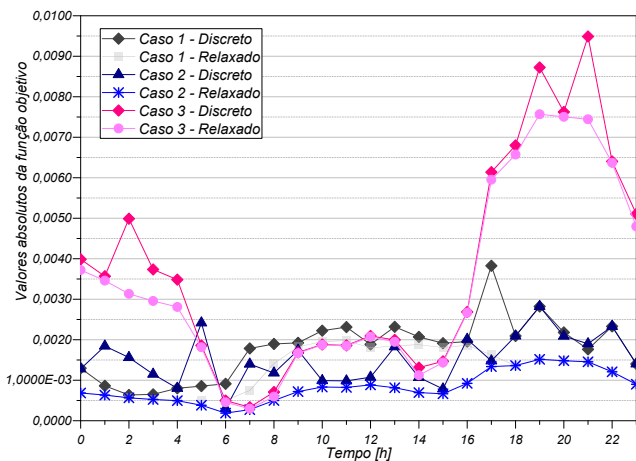


Figura 11: Valores da função objetivo para os casos avaliados do problema discreto e com relaxação contínua das variáveis de controle da RD de 135 nós.

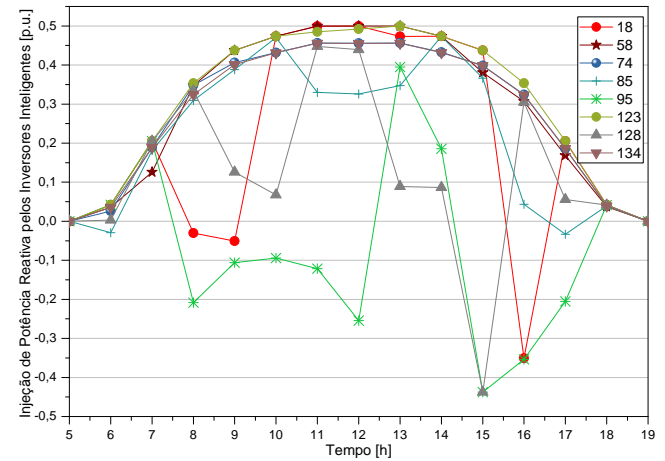


Figura 13: Perfis de injeção de potência reativa pelos inversores inteligentes da RD de 135 nós para o estudo de caso 3.

VI. CONCLUSÕES

Este trabalho propôs um modelo para o problema de planejamento de véspera da operação horária de RDs com GFD integrada por inversores inteligentes por meio da formulação de um problema de OVV com variáveis de decisão con-

resultados obtidos aos das propostas apresentadas em [10, 11] são inviáveis.

tínuas e discretas. A metodologia empregada para a resolução deste problema de OVV consiste na modificação do problema original em um problema equivalente somente com variáveis de decisão contínuas e no tratamento das variáveis de decisão discretas por funções penalidade senoidal quadráticas. Pelos resultados obtidos, pôde-se concluir que a penetração da GFD integrada às RDs por inversores inteligentes diminui efetivamente os desvios de tensão nos alimentadores primários ao mesmo tempo que proporciona, de maneira geral, uma menor demanda de potência reativa pela RD, o que resulta na ampliação das margens de carregamento do sistema elétrico como um todo.

Trabalhos futuros considerarão a limitação do número de ajustes nas variáveis de controle ao longo do horizonte de planejamento, a determinação de uma função objetivo que considere simultaneamente a minimização nos desvios de magnitude de tensão e a minimização das perdas na RD, e a característica estocástica da GFD para horizontes de planejamento mais longos.

REFERÊNCIAS

- [1] H. Villarroel-Gutiérrez, "The Argentine electrical sector and its trends toward renewable resources," *IEEE Latin America Transactions*, vol. 17, no. 10, pp. 1625–1636, 2019.
- [2] P. Benalcázar, J. Lara, and M. Samper, "Distributed photovoltaic generation in Ecuador: economic analysis and incentives mechanisms," *IEEE Latin America Transactions*, vol. 18, no. 3, pp. 564–572, 2020.
- [3] F. Katiraei and J. R. Aguero, "Studies for utility-scale photovoltaic distributed generation," *IEEE Power and Energy Magazine*, vol. 9, no. 3, pp. 62–71, 2011.
- [4] V. Vita, T. Alimardan, and L. Ekonomou, "The impact of distributed generation in the distribution networks' voltage profile and energy losses," in *2015 IEEE European Modelling Symposium (EMS)*, 2015, pp. 260–265.
- [5] J. Hou, Y. Xu, J. Liu, L. Xin, and W. Wei, "A multi-objective volt-var control strategy for distribution networks with high PV penetration," in *10th International Conference on Advances in Power System Control, Operation & Management (APSCOM 2015)*, 2015, pp. 1–6.
- [6] I. Lepadat, E. Helerea, S. Abagiu, and C. Mihai, "Impact of distributed generation on voltage profile and power losses in a test power grid," in *2017 International Conference on Optimization of Electrical and Electronic Equipment (OPTIM) & 2017 International Aegean Conference on Electrical Machines and Power Electronics (ACEMP)*, 2017, pp. 128–133.
- [7] G. Gutiérrez, A. Aguila, D. González, and L. Ortiz, "Optimum location and sizing of capacitor banks using VOLT VAR compensation in micro-grids," *IEEE Latin America Transactions*, vol. 18, no. 3, pp. 465–472, 2020.
- [8] B. Seal, *Common Functions for Smart Inverters*, 4th ed. Electric Power Research Institute, 2016.
- [9] A. Molina-García, R. A. Mastromauro, T. García-Sánchez, S. Pugliese, M. Liserre, and S. Stasi, "Reactive power flow control for PV inverters voltage support in LV distribution networks," *IEEE Transactions on Smart Grid*, vol. 8, no. 1, 2017.
- [10] F. M. de Vasconcelos, G. R. M. da Costa, and G. G. Lage, "An NLP penalty-based strategy for handling discrete controls for Volt/Var optimization in distribution systems," in *2015 IEEE Power & Energy Society General Meeting*, 2015, pp. 1–5.
- [11] M. B. Liu, C. A. Cañizares, and W. Huang, "Reactive power and voltage control in distribution systems with limited switching operations," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 889–899, 2009.
- [12] P. Belotti, C. Kirches, S. Leyffer, J. Linderoth, J. Luedtke, and A. Mahajan, "Mixed-integer nonlinear optimization," *Acta Numerica*, vol. 22, pp. 1–131, 2013.
- [13] E. M. Soler, V. A. de Sousa, and G. R. M. da Costa, "A modified primal-dual logarithmic-barrier method for solving the optimal power flow problem with discrete and continuous control variables," *European Journal of Operational Research*, vol. 222, no. 3, pp. 616–622, 2012.
- [14] E. M. Soler, E. N. Asada, and G. R. M. da Costa, "Penalty-based nonlinear solver for optimal reactive power dispatch with discrete controls," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 2174–2182, 2013.
- [15] L. Liu, X. Wang, X. Ding, and H. Chen, "A robust approach to optimal power flow with discrete variables," *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1182–1190, 2009.
- [16] Mingbo Liu, S. K. Tso, and Ying Cheng, "An extended nonlinear primal-dual interior-point algorithm for reactive-power optimization of large-scale power systems with discrete control variables," *IEEE Transactions on Power Systems*, vol. 17, no. 4, pp. 982–991, 2002.
- [17] F. Capitanescu and L. Wehenkel, "Sensitivity-based approaches for handling discrete variables in optimal power flow computations," *IEEE Transactions on Power Systems*, vol. 25, no. 4, pp. 1780–1789, 2010.
- [18] R.-H. Liang, Y.-K. Chen, and Y.-T. Chen, "Volt/Var control in a distribution system by a fuzzy optimization approach," *International Journal of Electrical Power & Energy Systems*, vol. 33, no. 2, pp. 278–287, 2011.
- [19] Z. Hu, X. Wang, H. Chen, and G. A. Taylor, "Damping subsynchronous resonance," *Engineering and Technology*, vol. 9, no. 4, pp. 431–442, 2002.
- [20] F. B. Prioste and G. R. Alvares, "Impactos da interligação de uma usina solar fotovoltaica com controle ativo de tensão em um sistema de distribuição," in *XIV Simpósio de Especialistas em Operação e Planejamento da Expansão Elétrica (SEPOPE)*, 2018, pp. 1–11.
- [21] Q. Long, J. Wang, D. Lubkeman, N. Lu, and P. Chen, "Volt-Var optimization of distribution systems for coordinating utility voltage control with smart inverters," in *2019 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, 2019, pp. 1–5.
- [22] C. Li, V. R. Disfani, H. V. Haghi, and J. Kleissl, "Coordination of oltc and smart inverters for optimal voltage regulation of unbalanced distribution networks," *Electric Power Systems Research*, vol. 187, p. 106498, 2020. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0378779620303011>
- [23] E. R. S. de Ferreira, R. M. Barros, T. A. R. d. Silva, R. A. L. de Rabêlo, V. R. Júnior, and G. G. Lage, "Application of a data communication infrastructure for the voltage magnitude control in transmission power systems," in *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, 2019, pp. 4308–4315.
- [24] V. Loia, A. Vaccaro, and K. Vaisakh, "A self-organizing architecture based on cooperative fuzzy agents for smart grid voltage control," *IEEE Transactions on Industrial Informatics*, vol. 9, no. 3, pp. 1415–1422, 2013.
- [25] M. Ibrahim and M. M. A. Salama, "Smart distribution system volt/var control using distributed intelligence and wireless communication," *IET Generation, Transmission Distribution*, vol. 9, no. 4, pp. 307–318, 2015.
- [26] S. S. Rangarajan, E. R. Collins, and J. C. Fox, "Efficacy of a smart photovoltaic inverter as a virtual detuner for mitigating network harmonic resonance in distribution systems," *Electric Power Systems Research*, vol. 171, pp. 175 – 184, 2019. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0378779619300562>
- [27] M. E. Baran and F. F. Wu, "Optimal capacitor placement on radial distribution systems," *IEEE Transactions on Power Delivery*, vol. 4, no. 1, pp. 725–734, 1989.
- [28] R. Kavasseri and C. Ababei, "REDS: REpository of Distribution Systems," <http://www.dejazzer.com/reds.html>, 2020.
- [29] E. N. M. Silva, A. B. Rodrigues, and M. da Guia da Silva, "Stochastic assessment of the impact of photovoltaic distributed generation on the power quality indices of distribution networks," *Electric Power Systems Research*, vol. 135, pp. 59–67, 2016.
- [30] AMPL, Optimization Inc., "AMPL," <http://www.ampl.com>, 2017.

- [31] R. Fourer, D. M. Gay, and B. W. Kernighan, *AMPL*, 2nd ed. Duxbury Press, 2002.
- [32] ARKI Consulting & Development A/S, "CONOPT," <http://www.conopt.com>, 2019.

Evaluation of a Sliding Window mechanism as Data Augmentation in Emotion Detection on Speech

Matheus Almeida Farias da Silva¹, Tiago da Silva Almeida¹ and Rafael Lima de Carvalho¹

¹ Federal University of Tocantins, Computer Science Department, Tocantins, Brazil

Reception date of the manuscript: 12/02/2021

Acceptance date of the manuscript: 01/04/2021

Publication date: 13/04/2021

Abstract— Emotion analysis is an important field of study, with many applications for security, financial, and politician. Despite being a subjective branch of study, emotion analysis can be simulated by Machine Learning algorithms that are trained for this purpose, through cataloged audio datasets, they can recognize patterns in these media that could be related to corresponding emotion. Neural Network Algorithms are able to work on the recognition of these emotions, with a focus only on audio, known as Speech Emotion Recognition (SER). Neural Network Algorithms generally obtain unequal averages of referring results such as recognition of emotions when applied to different audio datasets. This research evaluates a Data Augmentation method called Slide Window, which generates more data samples in order to increase the averages of classification rates. The method has been applied to three public datasets: EMO-DB, SAVEE, and RAVEDESS. The experiments have shown effectiveness in the increasing of the recognition rates of about to 11.95% on the EMO-DB base, 22.76% on SAVEE, and 18.82% on RAVEDESS when compared to other approaches in the literature.

Keywords— Speech Emotion Recognition, Voice processing, Machine Learning, Deep Neural Networks

I. INTRODUCTION

Speech Emotion Recognition (SER) is the process of inferring the emotion through machine learning using speech as data input [1]. Thus, this field includes algorithms to extract spectral features related to voice inputs [2]. In this way, the further step consists of the machine automation in order to find a model capable of identifying patterns using those features.

Towards investigating the SER problem, Jiang *et. al.* [1] proposed an algorithm called Parallelized Convolutional Recurrent Neural Network, (PCRNN), which is composed of two modules: A) CNN (Convolutional Neural Network) and B) LSTM (Long Short-Term Memory). Each audio sample is submitted to both modules, and the final scores are ensemble in order to produce a unique output. Module A is responsible to learn details of emotion using features such as speech frequency. On the other hand, module B tries to identify temporal changes in emotions. Their proposed work was submitted to four labeled audio dataset, known as EMO-DB, CASIA, ABC, and SAVEE. For which, they have calculated Weighted Average and Average on the data of confusion matrix.

In [3] an emotion recognition algorithm is developed with a 2D CNN architecture in which it intends to optimize train-

ing time and enhance the selection of key data, with the main focus on audio pre-processing. It means to focus its process on eliminating noises and intervals without any speech. Their experiments were carried out on the IEMOCAP and RAVEDESS datasets. Furthermore, it is mentioned that the focus of execution in the pre-processing was centralized by the fact that the current CNN architectures did not reveal any significant improvement in terms of precision and cost complexity in the processing of voice signals, and the use of LSTM is useful for training sequential data, but they are difficult to train effectively and are more computationally complex.

In the pre-processing stage of [3], the audio files undergo noise filtering and removal of irrelevant information using a function that takes the energy and amplitude data as input, producing a single value for the analyzed audio segment. After this stage, the spectrogram of these audios is used as a two-dimensional input representing the strength of the audio signal by the different frequencies.

Zhao *et. al.* [4] proposed two architectural models (1D CNN LSTM and 2D CNN LSTM) for recognition of emotions using voice. In order to increase the accuracy of emotion recognition, the authors proposed a learning block called Local Feature Learning Block (LFLB), which is responsible for the convolutional learning process. This block consists of a convolutional layer, a normalization layer, an exponential linear unit (ELU) layer, and a Max-Pooling layer, which is used to reduce the scales of the characteristic matrix. The algorithm contains three LFLBs, which generate a partial result to be inserted into an LSTM network (final classifier).

The 1D model uses the raw audio data as input, while the 2D uses the log of Mel scale as its input. The algorithm has been submitted under the German EMO-DB and the American IEMOCAP datasets.

Therefore, this work aimed to build a Neural Network based on [4] and evaluate the effectiveness of a Data Augmentation method called the Sliding Window under the EMO-DB, RAVDESS, and SAVEE datasets. The considered Sliding Window method consists of creating new audio files from the original files of each base. A fixed size window is applied to the audio, generating new audio files with the same size.

The paper is structured as follows. In Section II we describe the Mel-Frequency Cepstral Coefficients and the Deep Neural Network algorithms used in the experiments. In Section III we describe the Sliding Window mechanism for Data Augmentation as well as the benchmark datasets. In addition, in Section IV we describe the experiments and the results about the evaluated methodology. Finally, Section V brings the final remarks of the proposed work.

II. BACKGROUND

In this section we present the features and classifiers employed in this research.

a. Mel-Frequency Cepstral Coefficients

Psychophysical studies have shown that human perception of speech sound frequency does not follow a linear scale [5]. The Mel scale is a method that tries to reproduce that perception. It collects speech parameters similar to those used by humans to hear speech while considering all other information [6]. Thus, for each real frequency sonorous tone, f , measured in Hz, a concurrent tone is measured on a scale called the Mel scale f_{mel} ,

$$f_{mel} = 2595 \log_{10} \left(t1 + \frac{f}{700} \right), \quad (1)$$

which is a scale that aims to imitate the unique characteristics perceived by the human ear.

The normalization of frequencies in Mel scale consists of two main steps: i) the audio frames (small audio segment) are filtered in a process called Fast Fourier Transform (FFT), for which filters are applied with the average of the spectrum around the central frequency, and have different window sizes. These window sizes refer to the number of samples (of raw audio). In this way, the larger the window, the fewer filters are generated. This transformation process is called filter bank; ii) the resulting frequencies from step i) are normalized to the Mel scale [7].

In order to generate Mel-Frequency Cepstral Coefficients (MFCC), the first step is to divide the audio signal into frames, after that, they will be applied to the FFT. Second, filter bank processing is performed in the power spectrum, using the Mel scale. Obtaining the MFCC coefficients can be described as,

$$\hat{C}_n = \sum_{k=1}^k (\log \hat{S}_k) \cos \left[n \left(k - \frac{1}{2} \right) \frac{\pi}{k} \right] \quad (2)$$

where k is the number of Mel coefficients, \hat{S}_k is the output of the filter bank, and \hat{C}_n is the final MFCC coefficient.

b. Classifiers

1. Long Short-Term Memory

LSTM contains special units called memory blocks in the recurring hidden layer. The memory blocks contain memory cells with self-connections that store the temporal state of the network, in addition to special multiplicative units called gates to control the flow of information and assign of priorities [8].

2. Convolutional neural network

CNN has been used for pattern recognition tasks, such as facial recognition and handwritten numeric recognition [9]. CNN learns to map a particular image, in this case, a matrix of values, to its corresponding category, detecting a series of specific characteristics of each entry [10]. A CNN can have dozens or hundreds of layers, where each one learns to detect different characteristics of a matrix.

III. METHODOLOGY

This section presents the construction techniques and methods that were the basis for obtaining the results of the study proposed in this work. In subsection a it is shown the datasets used in the experiments, with basic information such as the amount of samples for each class (emotion). In addition, in subsection 2 we explain how the sliding window mechanism works. In the further subsections we describe the evaluation functions as well as the employed classifier.

a. Datasets

1. Surrey Audio-Visual Expressed Emotion

Surrey Audio-Visual Expressed Emotion (SAVEE) is an audiovisual dataset that was designed with the purpose of being applied in the development of automatic emotion recognition systems [11]. It consists of recordings from 4 male actors, graduate students, and researchers from the University of Surrey aged between 27 and 31, with a total of 7 emotions. Emotions are described psychologically in discrete categories: anger, disgust, fear, happiness, sadness, and surprise. Neutral emotion was also added to provide the 7 desired categories.

2. Ryerson Audio-Visual Database of Emotional Speech and Song

The Ryerson Audio-Visual Database of Emotional Speech and Song, called RAVDESS, is a Canadian multimodal dataset of speech and music with a focus on emotional analysis that has been entirely recorded in English. RAVDESS consists of 24 professional actors, each performing 104 unique vocalizations with emotions that include: happiness, sadness, anger, fear, surprise, disgust, calm and neutral [12]. Because it is multimodal, it is divided into three modules. In the first, audio-visual (AV), each file contains both the video and audio recording. The second, video-only (VO), consists

only of the video of actors facial expressions. And finally, the audio-only (AO), containing only the recorded audio.

3. Berlin Database of Emotional Speech

Containing speeches of 10 actors, 5 women and 5 men, the Berlin Database of Emotional Speech (EMO-DB) [13] is a German dataset of audios for the analysis of emotion. The actors generated 10 phrases each, half of which were long phrases, the other half, short phrases and some extra phrases, which can be used in daily communication and are interpretable in all applied emotions. Emotions are divided into 7 categories: neutral, anger, fear, joy, sadness, disgust, and boredom. The recordings were made in an anechoic chamber with high-quality recording equipment. The material contained in the dataset comprises about 800 phrases.

b. Data and Pre-Processing

1. MFCC

The library Librosa [14] has methods of extracting the audio characteristics, among them the MFCC, being able to choose the number of coefficients to be generated. Whereas, the greater the number of coefficients, the more detailed the data returned will be. Thus, in this work, 128 coefficients were used, with the FFT with a window size of 2048 frames per filter, and a shift of 512 frames per filter. In order to keep all audios at the same size, shorter audios have been extended up to 8 seconds, by concatenating their own content. On the other hand, audios greater than 8 seconds have been truncated to this standard size. Consequently, as all audios have 8 seconds, the resulting matrix will be 128x251 for all audios, with the first value being the number of coefficients and the second the number of resulting values.

2. Sliding Window Mechanism

Overfitting is a very recurring problem found in neural network training [15]. It makes the machine learning algorithm very good at recognizing the patterns of the training data but does not get as good at recognizing the test data. Overfitting limits the generalization of the algorithm. Some techniques used during the construction of the neural network algorithm can avoid this problem, one of which is Data Augmentation, which is related to the generation of new samples from the existing data, in order to increase the size of the dataset [16]. Thus, the neural network algorithm receives a larger amount of data, both for training and for testing, and consequently, its generalization capacity will be greater, since it will be able to recognize more data samples.

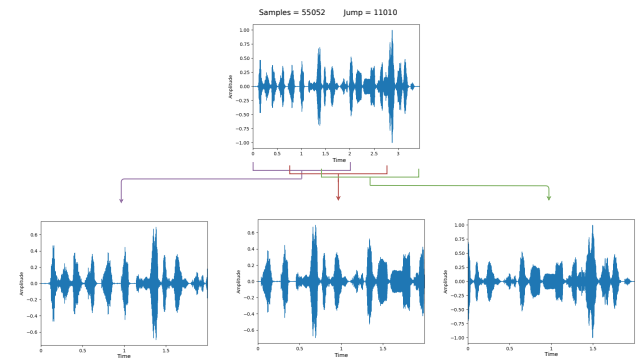
For this purpose, this work used a method called Sliding Window, as a way to perform the Data Augmentation technique. Once the window and shift sizes are defined, the window is positioned at the beginning of the audio data, cropping the file as it slides over it. Each cropped area is a new sample of the same class.

Window Structure

In order to process the audios, a sample rate is defined. The sample rate represents the number of frequency samples to be collected per second. This work uses a sample rate of

16.000 (sixteen thousand), thus, with a Sliding Window of 2 seconds, each sample is composed of 32.000 frequencies. In addition, audio files with less than 2.5 seconds are not suitable for running the sliding window mechanism, so they have been avoided by this procedure. An example where a single audio with 3.44 seconds became three new ones with 2 seconds each is shown in Fig. 1.

Fig. 1: Audio multiplication in three new audios through the Sliding Window process.



Slide Window Shift Size

As we can see in Section III, the audio files have different sizes. If we establish a fixed number for the shift of the window, this would cause an extremely unbalanced augmented dataset. In order to overcome this problem, the shift size is calculated using the equation 3.

$$S = \frac{tam}{Df} \quad (3)$$

where S is the amount of frequency data that the Sliding Window moves and tam the total size of frequency data for the audio input. The Df (coefficient of division) determines the size of the Sliding Window shift by dividing the number of frequency samples of the original audio. In this work, we defined $Df = 5$ as a good value based on the fact that the minimum accepted audio length (2.5 seconds) in the Sliding Window process and the size of the resulting audios (2 seconds). Establishing $Df = 5$ guarantees that at least two audio files (with 2 seconds each) will be generated for audio files with the minimum acceptable size.

Dataset transformation

The amount of audios generated from the Sliding Window process is directly related to the duration of each audio input, so the growth on the number of samples may vary for different emotions and datasets. If we take EMO-DB dataset we find audio samples that vary in size from 1.12 to 8+ seconds. In Table 1 it is shown the distribution of audio files by intervals of seconds. It can be seen in Table 1 that the RAVDESS dataset has practically all of its data over the 3 to 4 seconds category.

The augmentation data generated by the sliding window mechanism for each dataset and each emotion is shown in Table 2. We observe that the growing is different for each dataset. This occurs due to the duration of the audio samples. Samples with greater audio sizes produces more new

TABLE 1: DISTRIBUTION OF THE AUDIO DURATION OVER THE EMOTION DATASETS.

Audio length category	EMO-DB	RAVDESS	SAVEE
1 to 2 seconds	126	0	13
2 to 3 seconds	224	1	92
3 to 4 seconds	136	997	191
Greater than 4 seconds	49	250	184
TOTAL	535	1248	480

samples. Looking at the datasets RAVDESS and SAVEE, we notice that their audio samples have an average length of 3.42 and 3.49 seconds. This implied a growth that remained above 200%. In contrast, the EMO-DB, with 2.46 seconds of average, obtained a lower growth, for which only half of its audios had been applied to the Sliding Window process.

TABLE 2: DATA GROWTH AFTER THE SLIDING WINDOW PROCESS.

EMO-DB			
Emotion	Original	With Window	Growth
Anger	127	287	126%
Boredom	81	195	141%
Disgust	46	141	207%
Fear	69	119	72%
Happiness	71	150	111%
Neutral	79	143	81%
Sadness	62	223	260%
TOTAL	535	1258	135%

RAVDESS			
Emotion	Original	With Window	Growth
Neutral	96	364	279%
Calm	192	759	295%
Happiness	192	748	290%
Sadness	192	755	293%
Anger	192	762	297%
Fear	192	728	279%
Disgust	192	764	298%
TOTAL	1248	4880	291%

SAVEE			
Emotion	Original	With window	Growth
Anger	60	209	248%
Disgust	60	224	273%
Fear	60	211	252%
Happiness	60	221	268%
Neutral	120	422	252%
Sadness	60	245	308%
Surprise	60	218	263%
TOTAL	480	1750	264.5%

c. Training methods

In this work, we consider the Speaker-Dependent (SD) approach to considering the training of the machine learning algorithm. In this method, the dataset is split into two parts: training and test. When separating the data, the algorithm is trained using the training data set and the evaluation is done using the test set. In the context of Voice Recognition, the name Speaker-Dependent means that the algorithm depends on the data of all Actors contained in the dataset, both in the test stage and in the training. In this work, the use of the Speaker-Dependent method considering 80% of data for training and the remaining 20% for tests, with stratification.

1. Evaluation functions

Unweighted Average

The Unweighted Average (UA) is an arithmetic average based only on the true positives, disregarding the other possible results obtained by the classes. This way of evaluating the algorithm performance with the Confusion Matrix is presented in works related to SER, as in [4] and in [1]. This work uses this evaluation method as the main reference, because it is one of the most used, in addition to being a great metric when the data are balanced.

ROC-AUC

ROC-AUC is also another way to assess the sensitivity of the algorithm in classifying test data. ROC (Receiver Operating Characteristic) is represented by a graphical curve generated from a function that calculates the true positives with the false positives. In this case, a unique number (score) named AUC (Area Under the Curve) shows the area that the curve covers (used to evaluate the algorithm). The higher the ROC-AUC, the greater the ability of the algorithm to distinguish between classes.

F1-Score

Divided into Macro and Micro, the F1-Score is an evaluation metric on the test data that aims to show the degree of precision the algorithm has, taking into account the balance of each class. In this sense, the more similar the algorithm precision (in the case of this work, the UA), the more the data is balanced, proving its precision veracity. For the F1-Score with the Micro approach, the F1-Score measures a single value taking into account the balance of the classes as a whole, whereas the Macro, originates from the calculation with each class.

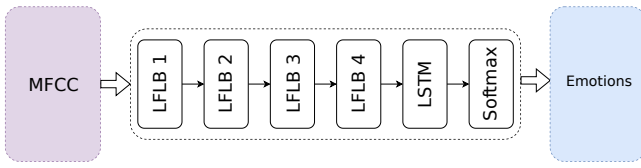
d. Classification Algorithm

Making a computer recognize an emotion requires, mostly, a deep and broad analysis in relation to the audio characteristics. It is important to specify how each value consists, as well as the relationship of these values as a whole over time. With this in mind, this work used the same algorithm model present in [4]. The model is composed of four Local Feature Learning Blocks (LFLB), which are blocks composed of layers with focused CNN 2D to abstract as much relevant information from the audio as possible, and an LSTM network layer. This scheme has been selected with the objective of analyzing dependencies between a sequence of data. Its main objective is to identify which are the dependencies of temporal context from the data resulting from the LFLB, and finally a Softmax layer to the generalization of logistic regression for classification problems with several classes. Those classes are the emotions of the audio according to the learned characteristics. In Fig. 2 it is shown the arrangement of the algorithm used for this work.

1. CNN 2D

In order to obtain a higher level of abstraction of which audio characteristics result in the classification of a certain emo-

Fig. 2: Architecture diagram of the algorithm for the classification of emotions from the MFCC extracted from the audios [4].



tion, the layers of the CNN 2D neural network were used for the composition of the LFLBs. These types of layers rely on the most outstanding data filtering, which is done through a kernel. For all convolution layers, a 3x3 size kernel is employed, that is, 9 units of the characteristics matrix passed through the filtering at a time, with a 1x1 Stride, which informs how many units the kernel has moved until its next filtering step. A small number of Stride is also used in order to capture more information about the audio input.

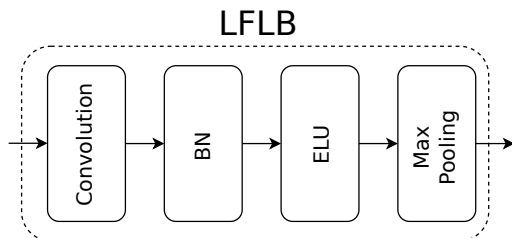
2. LSTM

The LSTM layer in this algorithm has the purpose of examining the composition and variation of the audio data coming from the CNN layers through a temporal analysis. It is responsible for generating the latest classification data of the neural network. Each cell (memory block) composing the LSTM layer performs an analysis calculation on the inserted data, thus, a total number of 256 LSTM cells is used, twice the output value of the last CNN layer. This is necessary to increase the reliability of the generated result.

3. Local Feature Learning Block

Each LFLB is composed of a Convolutional layer, a Batch Normalization (BN), an ELU layer (exponential linear unit), and a Max Pooling layer, with the Convolutional and Max Pooling layers being the fundamentals of this set [4]. Thus, the composition of an LFLB and the corresponding operation order are shown in Fig. 3.

Fig. 3: Composition of a Local Feature Learning Block [4].



As mentioned in Subsection 2, the Convolutional layers filter the input data as a way of mapping the striking features. These features are also called activations. The BN layer normalizes the results of the Convolutional layer of each batch of trained audios, this implies an improvement in performance and stability of the learning layers. The transformation applied by BN keeps the average activation close to 0 and the standard deviation of activation close to 1. The BN layer is then connected to the ELU that calculated the BN output. It takes the average of the activation values towards zero, in order to decrease the learning time and increase the level of

recognition. Finally, the Max Pooling layer takes the most important features to stand out over the distortions and noise contained in the audios. It manages to make this abstraction by dividing the input into regions and giving the maximum values for each region as outputs.

4. Layer parameters

In order to make sure the correct execution of the classification algorithm, it is important to pay attention to some implementation details. One of these details is the control of data produced from each layer. The size of the output of each LFLB is defined by the kernel size, i.e. the region that filters the activations of the layer, and the Stride that tells how many units the kernel must move after activation. And for the LSTM layer, the size of its output is defined by how many units of LSTM the layer is composed of. Table 3 presents the control configuration used in this work.

IV. RESULTS

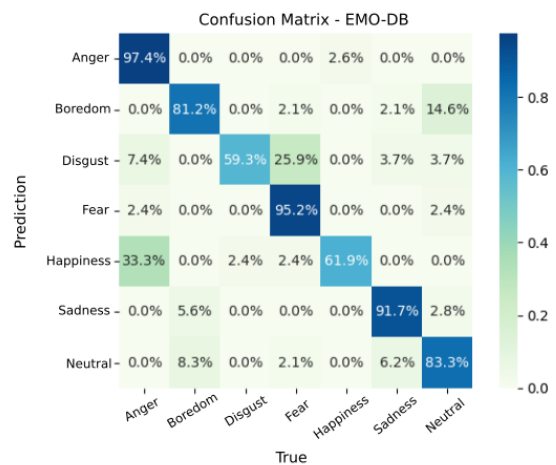
a. Test configuration

Audio files have been split using stratification and considering the amount of 80% for training and 20% for the test. In order to prevent overfitting, the mechanism of Early Stopping is employed on training data. Empirically, was defined an amount of 15 epochs as a marker. After this number, if the neural network does not improve accuracy under the validation set, the training is stopped and the best configuration is saved. Each emotion dataset has been trained 3 times. We averaged the following metrics: confusion matrix, UA, ROC-AUC, Micro, and F1-Scores.

b. Results without Sliding Window

We organized the experiments by first presenting the results of the classification algorithm without the proposed data augmentation mechanism. The objective here is to show the neural network generalization capabilities and general performance using the MFCC features as input.

Fig. 4: Confusion matrix resulting from training the EMO-DB dataset without the Sliding Window.

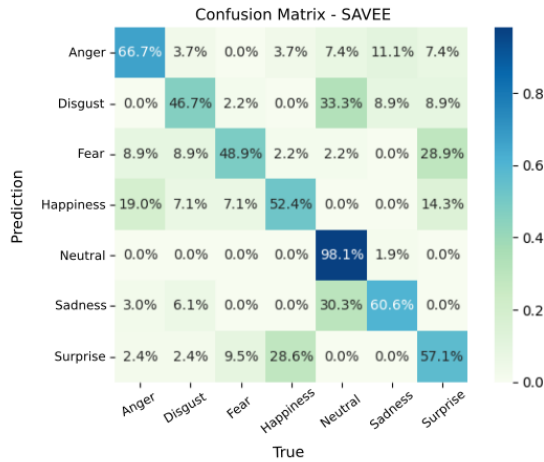


In Fig. 4 it is shown the confusion matrix of the classifier

TABLE 3: SET OF PARAMETERS USED IN THE CLASSIFICATION ALGORITHM.

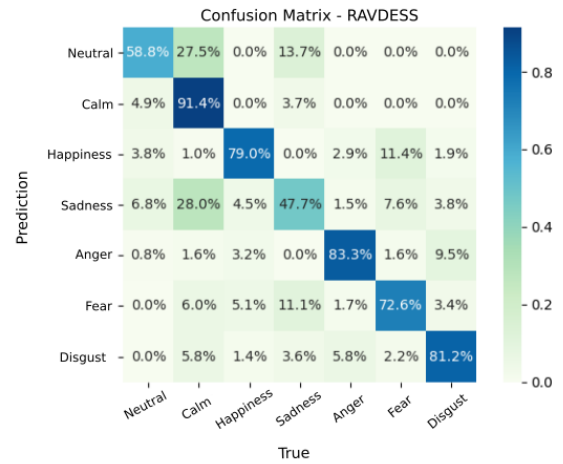
Layer	Output size	Kernel Size	Stride
LFLB 1 - Convolution	128 x 256 x 64	3 x 3	1 x 1
LFLB 1 - Max Pooling	64 x 125 x 64	2x2	2x2
LFLB 2 - Convolution	64 x 125 x 64	3 x 3	1 x 1
LFLB 2 - Max Pooling	16 x 31 x 64	4 x 4	4 x 4
LFLB 3 - Convolution	16 x 31 x 128	3 x 3	1 x 1
LFLB 3 - Max Pooling	4 x 7 x 128	4 x 4	4 x 4
LFLB 4 - Convolution	4 x 7 x 128	3 x 3	1 x 1
LFLB 4 - Max Pooling	1 x 1 x 128	4 x 4	4 x 4
LSTM	256	-	-
Softmax - Dense	7	-	-

using the EMO-DB. It can be seen that some emotions such as Fear and Sadness achieved a high percentage of recognition, even with a lower amount of data analyzed than the others belonging to the EMO-DB base. Although other emotions such as Disgust and Happiness, obtained much lower results, this behavior might be explained by the lack of sufficient data for the algorithm to be able to precisely recognize the differences in such emotions. As a general result, we observe that Happiness maintains a low percentage of recognition in all datasets, as can be seen in Figs.4, 5 and 6. In addition to the number of audio files, what causes this occurrence is the fact that the algorithm is often confused with more intense emotions, such as Anger and Surprise for example.

Fig. 5: Confusion matrix resulting from training the SAVEE dataset without the Sliding Window.

Observing the experiments without the Sliding Window, it is important to observe the performance variation involving the same emotion throughout different datasets. An interesting case occurs with emotion Disgust shown in Figs. 5 and 6. Even having the same number of samples as the other emotions for both datasets (as can be seen in Table 2), the discrepancy of performance values is still large.

Table 4 presents the final results of the experiments without the sliding window mechanism. It can be observed a difference of practically 20% of the recognition between the EMO-DB and SAVEE datasets, when comparing the UA function. The dataset SAVEE, which has the smallest number of audio sample, obtained the worst results in training.

Fig. 6: Confusion matrix resulting from training the RAVDESS dataset without the Sliding Window.

c. Results with Sliding Window

Figs. 7, 8 and 9, present the appropriate confusion matrices generated by the experiments in each dataset. In contrast to the confusion matrices in Subsection b, it appears that there has been a reduction in the disparity of results between emotions, such as the disgusting and neutral emotions of the SAVEE dataset in Fig. 5 there is a difference of recognition of 51.4 %. However, the corresponding results with the Sliding Window in Fig. 8 show a much smaller difference, with a percentage of 17.7 %. This evolution is due to the increase in audio samples that the algorithm produced, this made it better to calculate the weights of its layers, thus leading to a better classification capacity.

Due to the increase in data samples being different in each dataset, it implied distinct factors of growth in the final emotions recognition. In Table 5 it is shown the results obtained with the Sliding Window with the evaluated performance functions, having UA as the main observation parameter, and in Table 2 the recognition increase in each dataset is noticeable. In EMO-DB fewer audios were generated than the other datasets (as can be seen in Table 1), which led to a smaller increase in its final performance.

The promising results obtained with the Sliding Window mechanism, reinforced the idea that the increase in the amount of audio samples had led to the improvement of the algorithm's capacity to classify the input emotions.

TABLE 4: EVALUATION FUNCTIONS GENERATED FROM THE EXPERIMENTS WITHOUT THE SLIDING WINDOW.

Model	Input	Dataset	ROC-AUC	F1-Score Macro	F1-Score Micro	UA
Without SW	MFCC	EMO-DB	98.19%	82.50%	84.50%	81.44%
Without SW	MFCC	SAVEE	90.34%	61.1%	62.5%	61.50%
Without SW	MFCC	RAVDESS	95.29%	72.21%	73.60%	73.44%

SW = Sliding Window.

Fig. 7: Confusion matrix resulting from training the EMO-DB dataset with the Sliding Window.

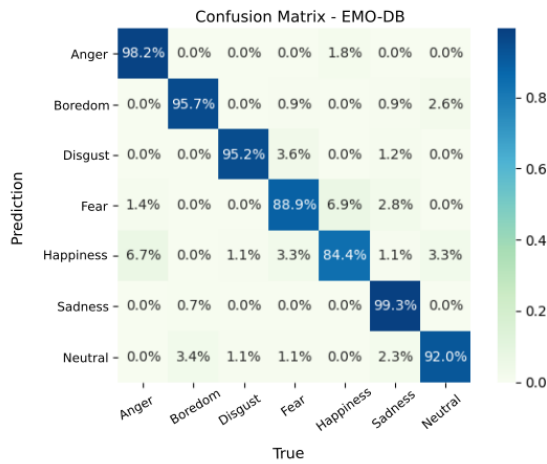


Fig. 9: Confusion matrix resulting from training the RAVDESS dataset with the Sliding Window.

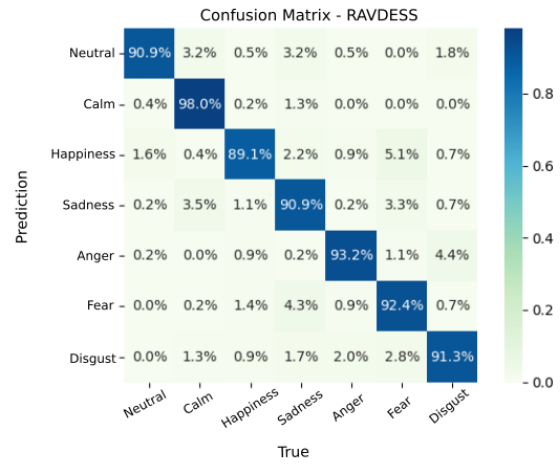


Fig. 8: Confusion matrix resulting from training the SAVEE dataset with the Sliding Window.

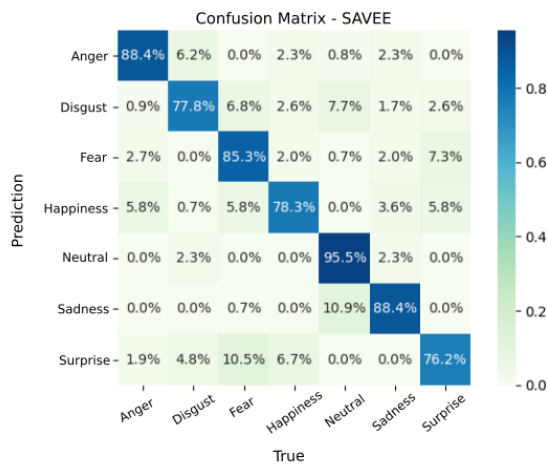


TABLE 5: EVALUATION FUNCTIONS GENERATED FROM THE EXECUTION OF THE DATASETS WITH THE SLIDING WINDOW.

Model	Input	dataset	ROC-AUC	F1-Score Macro	F1-Score Micro	UA
With SW	MFCC	EMO-DB	99.70%	93.64%	94.44%	93.9%
With SW	MFCC	SAVEE	98.21%	84.61%	86.00%	84.26%
With SW	MFCC	RAVDESS	99.52%	92.39%	92.38%	92.26%

SW = Sliding Window.

d. Comparison with related works

Tables 7, 8 and 9, show the comparison of the results between external works and the proposed methodology. In [1], the authors trained the classification algorithm with the audios of all the actors, except one. Therefore, the audios of this selected actor are not included in the training, and are used only in the test. This approach is known as LOSO (Leave

one Speaker Out).

TABLE 7: COMPARISON OF EMO-DB RESULTS WITH RELATED WORKS

Work	Training method	UA
[1]	LOSO	84.53%
[4]	Speaker Dependent	95.02%
Proposed	Speaker Dependent	93.39%

TABLE 8: COMPARISON OF SAVEE RESULTS WITH RELATED WORKS

Work	Training method	UA
[1]	LOSO	59.40%
Proposed	Speaker Dependent	84.26%

TABLE 9: COMPARISON OF RAVDESS RESULTS WITH RELATED WORKS *THIS IS THE RESULT DISREGARDING THE SURPRISE EMOTION, PRESENT IN [3] BUT NOT IN THIS WORK.

Work	Training method	UA
[3]	Speaker Dependent	*76.85%
Proposed	Speaker Dependent	92.26%

The results generated by this work managed to obtain an excellent performance in relation to the aforementioned works, except for [4]. As can be seen in Table 7, the work in question still managed almost 2% more in the UA assessment.

V. FINAL REMARKS

Simulating the process of recognizing human emotions on a computer, as already mentioned, is a very complex process. In addition to each emotion, there are cultural variations that can lead to different intonations. Such differences

may be very hard to capture during the classification process. This work presented a method of data augmentation of the audio samples called the Sliding Window. The algorithm used by this work, together with the Sliding Window method as pre-processing, managed to increase the accuracy of the emotions recognition in the EMO-DB dataset by 11.95%, in SAVEE by 22.76%, and RAVDESS at 18.82%. The results also have shown that the proposed method is competitive with state-of-the art approaches. The future investigations include the exploration of the window size parameter in order to determine the sensibility of the classifier over this parameter.

Having a consistent emotion recognition algorithm, which can discern the emotion of incoming audio, can be very useful in daily tasks and in more specific applications. Bearing this in mind, the application of the Sliding Window or similar method may be useful for future works hopping to achieve even more precise classification results.

REFERENCES

- [1] P. Jiang, H. Fu, H. Tao, and P. L. L. Zhao, "Parallelized convolutional recurrent neural network with spectral features for speech emotion recognition," *IEEE Access*, vol. 7, 2019. [Online]. Available: <https://ieeexplore.ieee.org/document/8756261>
- [2] Z. W. Fen Xu, "Emotion recognition research based on integration of facial expression and voice," *2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, vol. 7, 2018. [Online]. Available: <https://ieeexplore.ieee.org/document/8633129>
- [3] S. K. Mustaqeem, "A cnn-assisted enhanced audio signal processing for speech emotion recognition," *Sensors*, vol. 20, p. 183, 12 2019.
- [4] J. Zhao, X. Mao, and L. Chen, "Speech emotion recognition using deep 1d & 2d cnn lstm networks," *Biomedical Signal Processing and Control*, vol. 47, pp. 312 – 323, 2019. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1746809418302337>
- [5] M. A. Hossan, S. Memon, and M. A. Gregory, "A novel approach for mfcc feature extraction," in *2010 4th International Conference on Signal Processing and Communication Systems*, 2010, pp. 1–5.
- [6] N. Dave, "Feature extraction methods lpc , plp and mfcc in speech recognition," 2013.
- [7] S. A. Alim and N. K. A. Rashid, "Some commonly used speech feature extraction algorithms," in *From Natural to Artificial Intelligence*, R. Lopez-Ruiz, Ed. Rijeka: IntechOpen, 2018, ch. 1. [Online]. Available: <https://doi.org/10.5772/intechopen.80419>
- [8] H. Sak, A. Senior, and F. Beaufays, "Long short-term memory recurrent neural network architectures for large scale acoustic modeling," *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, pp. 338–342, 01 2014.
- [9] M. Matsugu, K. Mori, Y. Mitari, and Y. Kaneda, "Subject independent facial expression recognition with robust face detection using a convolutional neural network," *Neural Networks*, vol. 16, no. 5, pp. 555 – 559, 2003, advances in Neural Networks Research: IJCNN '03. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0893608003001151>
- [10] S. Khan, H. Rahmani, S. A. A. Shah, M. Bennamoun, G. Medioni, and S. Dickinson, *A Guide to Convolutional Neural Networks for Computer Vision*, 2018.
- [11] P. Jackson and S. ul haq, "Surrey audio-visual expressed emotion (savee) database," 04 2011.
- [12] S. R. Livingstone and F. A. Russo, "The ryerson audio-visual database of emotional speech and song (ravdess): A dynamic, multimodal set of facial and vocal expressions in north american english," *PLOS ONE*, vol. 13, no. 5, pp. 1–35, 05 2018. [Online]. Available: <https://doi.org/10.1371/journal.pone.0196391>
- [13] F. Burkhardt, A. Paeschke, M. Rolfes, W. Sendlmeier, and B. Weiss, "A database of german emotional speech," vol. 5, 01 2005, pp. 1517–1520.
- [14] Brian McFee, Colin Raffel, Dawen Liang, Daniel P.W. Ellis, Matt McVicar, Eric Battenberg, and Oriol Nieto, "librosa: Audio and Music Signal Analysis in Python," pp. 18 – 24, 2015.
- [15] L. Rice, E. Wong, and J. Kolter, "Overfitting in adversarially robust deep learning," 02 2020.
- [16] E. Lashgari, D. Liang, and U. Maoz, "Data augmentation for deep-learning-based electroencephalography," *Journal of Neuroscience Methods*, vol. 346, p. 108885, 2020. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0165027020303083>