UNIVERSIDADE DE BRASÍLIA
INSTITUTO DE CIÊNCIAS HUMANAS
DEPARTAMENTO DE GEOGRAFIA
PROGRAMA DE PÓS-GRADUAÇÃO EM GEOGRAFIA

DETECÇÃO DE DEPRESSÕES CÁRSTICAS NO BRASIL USANDO SEGMENTAÇÃO SEMÂNTICA E DE INSTÂNCIA E COMPARANDO DIFERENTES MODELOS DIGITAIS GLOBAIS DE ELEVAÇÃO.

HEITOR DA ROCHA NUNES DE CASTRO

Brasília

Junho de 2023

HEITOR DA ROCHA NUNES DE CASTRO

DETECÇÃO DE DEPRESSÕES CÁRSTICAS NO BRASIL USANDO SEGMENTAÇÃO SEMÂNTICA E DE INSTÂNCIA E COMPARANDO DIFERENTES MODELOS DIGITAIS GLOBAIS DE ELEVAÇÃO.

Dissertação de Mestrado submetida ao Departamento de Geografia da Universidade de Brasília, como parte dos requisitos necessários para a obtenção do grau de Mestre em Geografia, área de concentração Gestão Ambiental e Territorial (Linha de Pesquisa - Análise dos Sistemas Naturais).

Aprovada por:

_____

Osmar Abílio de Carvalho Júnior, Doutor (GEA-UnB)
(Orientador)

_____

Antonio José Teixeira Guerra, Doutor (UFRJ)
(Examinador externo)

_____

Claudinei Taborda da Silveira, Doutor (UFPR)
(Examinador externo)

Brasília-DF, 16 de junho de 2023

FICHA CATALOGRÁFICA

CASTRO, HEITOR DA ROCHA NUNES DE.

DETECÇÃO DE DEPRESSÕES CÁRSTICAS NO BRASIL USANDO SEGMENTAÇÃO SEMÂNTICA E DE INSTÂNCIA E COMPARANDO DIFERENTES MODELOS DIGITAIS GLOBAIS DE ELEVAÇÃO. / Heitor da Rocha Nunes de Castro; Orientador: Prof. Dr. Osmar Abílio de Carvalho Júnior. – Brasília, 2023.

60 f.: il. color.; 30 cm.

Dissertação (Mestrado em Geografia) - Universidade de Brasília, 2023.

1. Sinkholes. 2. Dolines. 3. Artificial Intelligence. 4 Deep Learning. 5. Karst Geomorphology. 6. Topographic Depression. 7. Geomorphometric Attributes I. Carvalho Jr., O. A., orient. II. Título.

AGRADECIMENTOS

Ao meu companheiro, Gilvan, pelo apoio incondicional em mais uma jornada de minha vida profissional e acadêmica;
Ao meu orientador, professor Dr. Osmar Abílio de Carvalho Jr., pelo aprendizado que me foi oportunizado, pelas sugestões, conversas e tranquilidade na orientação da pesquisa;
Aos professores Dr. Antonio José Teixeira Guerra e Dr. Claudinei Taborda da Silveira, que gentilmente aceitaram participar de minha Banca de Defesa;
À Maria Damasceno e Anesmar Albuquerque, que me ajudaram com muita presteza em minhas dúvidas sobre geotecnologias;
Aos demais professores e membros do LSIE e do Departamento de Geografia;
Aos amigos e colegas do IBAMA, por todo apoio;
Aos meus pais e irmãos;
Ao Eterno.

APRESENTAÇÃO

O carste carbonático (convencional) é uma paisagem complexa, em que a erosão química causada pelas águas das chuvas incide sobre rochas suscetíveis à ação hídrica. Ao longo do tempo geológico, essa interação permite que a paisagem cárstica exiba morfologias diversas, positivas e negativas, no exocarste e no endocarste, propiciando o aparecimento de diversas formas e processos típicos desse sistema.

As depressões, como as dolinas, estão entre as feições mais características do carste, podendo apresentar formas arredondadas ou angulares e de diversos tamanhos. São formadas a partir do colapso ou subsidência de terrenos sujeitos à dissolução e erosão pela ação química das águas. No Grupo Bambuí, particularmente na região meio-oeste da Bahia, as dolinas se encontram espalhadas na paisagem, com padrões de distribuição e densidade que indicam áreas potencialmente sensíveis aos impactos causados pelas atividades humanas.

No sentido de detectar as depressões cársticas da área de estudo, inserida no Grupo Bambuí, meio-oeste da Bahia, a pesquisa buscou identificar essas feições a partir de ferramentas de sensoriamento remoto, com método de aprendizado profundo (*deep learning*), que está no contexto da inteligência artificial, sem que houvesse, necessariamente, deslocamento para estudo de campo.

Primeiramente, foram obtidos dados de Modelos Digitais de Elevação (MDEs) das últimas versões dos sensores GLO-30, AW3D30, SRTM, NASADEM e ASTER GDEM, com resolução espacial de 30 metros, disponíveis gratuitamente. Aos MDEs, foram adicionadas variáveis que subdividiram os MDEs em dois tipos: MDE + MDE *sink-depth* (2 variáveis), e MDE + MDE *sink-depth* + 9 morfometrias - declividade, aspecto, perfil de convexidade, plano de convexidade, convexidade longitudinal, convexidade transversal, mínima curvatura, máxima curvatura e raiz quadrada média - (11 variáveis).

Para a interpretação visual das feições (depressões cársticas), foram obtidas imagens ópticas de dados Sentinel-2 e OLI-Landsat 8. Os vetores foram criados, editados e manipulados, gerando um arquivo vetorial mostrando a verdade do terreno.

Em seguida, foram aplicados métodos de aprendizado profundo (*deep learning*), em que foram utilizadas cinco arquiteturas de segmentação semântica (FPN, LinkNet, Unet, Unet ++ e DLV3+) e uma de segmentação de instâncias (Mask-RCNN). Na segmentação semântica, foi feita a análise de acurácia por *pixel* (*Overall Accuracy*,

Precision, *Recall*, *F-Score* e *IoU*), e na segmentação de instâncias foi feita a análise de acurácia por *Average Precision* (AP, AP50 e AP75).

Além disso, houve reconstrução da imagem de segmentação semântica (GLO-30 com 11 variáveis e arquitetura FPN), com janelas deslizantes em amostras de 128x128 *pixels*; reconstrução da imagem de segmentação de instâncias, com cinco limites (*Threshold Values*) considerando a acurácia por polígono (70%, 75%, 80%, 85% e 90%); e conversão semântica-instância.

Nos resultados, a segmentação semântica, com utilização da arquitetura FPN e do MDE GLO-30 (11 variáveis) obteve o melhor resultado (IoU = 71.77), e a conversão semântica-instâncias teve alto desempenho, em razão da ausência de sobreposição ou interação entre as feições. A pesquisa gerou uma base de dados numerosa de detecção das depressões cársticas da área de estudo, dado que a classificação com métodos de *deep learning* requer volume numeroso de dados, e com indicação dos resultados de classificação mais bem posicionados.

A discussão da pesquisa trata dos obstáculos encontrados na definição dos limites das dolinas, da comparação entre os cinco MDEs utilizados, da comparação entre as cinco arquiteturas de segmentação semântica, da segmentação de instâncias, da conversão semântica-instâncias, e das implicações do mapeamento de dolinas na análise ambiental da área de estudo.

A pesquisa buscou inovar ao aplicar métodos de *deep learning* na detecção de depressões cársticas por meio da utilização de MDEs de 30 m disponíveis gratuitamente, em área do Grupo Bambuí que apresenta relevância ambiental e está sujeita aos impactos de atividades humanas. Pesquisas futuras de aplicação de métodos semelhantes no carste do Grupo Bambuí, ou em outras paisagens cársticas, poderão ser realizadas, a partir de dados espaciais óticos, radar ou LiDAR.

A dissertação de Mestrado foi escrita na forma de artigo científico, que foi submetido à Revista Geomorphology.

ABSTRACT

This research aims to investigate the use of deep segmentation in detecting and quantifying natural karst depressions developed in the carbonate rocks of the Neoproterozoic Bambuí Group in Western Bahia, Brazil. The karst landscape of the study area has dolines and the formation of lakes enclosed in limestone. The study analyzes different approaches to detecting karst depressions. First, a comparison of five different Global Digital Terrain Models (DEM) with 30 meters resolution: Copernicus 30m Global DEM (GLO-30), ALOS World 3D (AW3D30), Shuttle Radar Topography Mission (SRTM), National Aeronautics and Space Administration DEM (NASADEM), and Advanced Spaceborne Thermal Emission and Reflection Radiometer - Global DEM (ASTER-GDEM). Second, comparing five semantic segmentation architectures with EfficientNet-B7 backbone (Feature Pyramid Network - FPN, LinkNet, Unet, Unet++, and DVL3+) and one instance segmentation (Mask-RCNN). Third, evaluation of segmentation elaboration using two variables (DEM and DEM-based sink depth) or eleven variables (DEM, DEM-based sink depth, and nine terrain attributes). The research did not evaluate the use of DEM in isolation due to its very low accuracy in previous analyses. The methodology had the following steps: (a) acquisition of DEMs and generation of geomorphometric attributes; (b) sample labeling by manual interpretation of karst depressions from Sentinel-2 and OLI-Landsat 8 images; (c) selection of samples for training (1600 samples), validation (400 samples) and testing (400 samples) with dimensions 128x128 considering two channels (DEM and depth of sinking based on DEM) and eleven channels (the two previous ones added by nine morphometric attributes); (e) elaboration of semantic and instance segmentations; (f) accuracy analysis; (g) image reconstruction using sliding window; and (f) conversion from semantic segmentation to instance using GIS tools. The results show that the GLO-30 data showed the highest accuracy values, followed by the AW3D30. In contrast, the ASTER GDEM obtained the worst results. Among the models using semantic segmentation, the FPN presented the most significant accuracy results, while the DVL3+ presented the worst. Considering the same architectures and DEM, the models that used 11 channels obtained better results than those that used only two channels. Converting data from semantic segmentation to instance segmentation using a GIS tool proved to be very easy since the features did not interact.

Keywords: Sinkholes, Dolines, Artificial Intelligence, Deep Learning, Karst Geomorphology, Topographic Depression, Geomorphometric Attributes.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Karst is a landscape produced by the chemical dissolution of water over soluble carbonate rocks such as limestone, dolomites, or gypsum, where chemical erosion processes overcome mechanical ones and lead to a distinct morphological and hydrological pattern. The carbonate rocks on which the karst model develops (outside glaciated areas) represent 15.2% of the continental land surface (Goldscheider et al., 2020). In South America, karsts cover 2% (370,809 km^2) of the continental surface, where most are in the Brazilian territory (containing 5-7% of the territory) spread across different phytogeographic domains (Bystriakova et al., 2019). Water is the main shaping agent of karst forms, leading to the disintegration and discontinuity of the rocky body subdivided into (a) exokarst, which involves positive surface morphologies (massifs, towers, and mogotes) and negative (sinkholes, gorges, and poljes); and (b) endokarst, consisting of underground features formed by water flow ducts, underground rivers, and caves.

Karst environments have various ecosystem services (Beynen, 2011; Fleury, 2009) and host a variety of ecological niches for animals and plants, including several endemic species (Bystriakova et al., 2019). Developing a specialized flora comes from environmental characteristics with high Ca, Mg, and K concentrations, absence of surface water, and low soil formation rates. These environments are fragile and susceptible to environmental risks from human activities, such as irrigation, pipe leaks, excavations, deforestation, artificial vibration, and heavy vehicle traffic (Gutiérrez et al., 2014). In addition, karst hydrological systems are highly vulnerable to polluting agents due to the rapid transfer of water flow between surface drainage and subsurface recharge areas (De Waele et al., 2011; Parise et al., 2015).

The primary morphological representation in the karst is the dolines, characterized by closed depressions, which vary in size (from a few to hundreds of meters) and with internal drainage. Dolines formation includes mechanisms of dissolution, collapse, suffusion, and regional subsidence (Ford and Williams, 2007). While the subsidence occurs slowly, the collapse happens suddenly, both induced by the physical-chemical erosion of the underground rock that sinks the overlying materials (Lee et al., 2016). The junction of two or more dolines forms uvulas, larger depressions with an irregular shape. Dolines generation is a natural process, but anthropic activities can induce their development due to changes in land use, urban growth, and implementation of

engineering projects that result in hydrological alteration of the karst (De Waele et al., 2011; Parise et al., 2015).

Mapping areas of karst depressions is a priority step due to the complexity and fragility of karst systems when exposed to anthropic modifications (Masilela and Beckedahl, 2022; Parise et al., 2015). Therefore, knowledge of the distribution of karst depressions is essential for planning land occupation, environmental protection, and management of sensitive and highly relevant areas (Hofierka et al., 2018). In addition, the delimitation of sinks helps to prevent fatalities in human lives and damage to properties (Lee et al., 2016). The unexpected presence of sinkholes, mainly in mineral extraction, human settlements, water discharges, and agriculture, increases the need to detect these features continuously. Fieldwork can provide accurate mappings, but they are expensive and time-consuming. Therefore, different methods have been proposed for the automated detection of karst depressions from Digital Elevation Models (DEM) and active and passive remote sensing images, which can be subdivided into traditional methods and methods based on Deep Learning (DL). Among the traditional methods, the following procedures stand out: (a) DEM-based sink depth (Antonić et al., 2001; Guimarães et al., 2005); (b) Topographic Position Index using a moving window operation that calculates the relative position between the elevation at each pixel in the DEM and the average elevation of the surrounding neighborhood (Doctor and Young, 2013; Kobal et al., 2015; Obu and Podobnikar, 2013; Weishampel et al., 2011); (c) delineation of depressions based on the Outermost Closed Contour (OCC) line using vector-based contour representation (Angel et al., 2004; Seale et al., 2008) and its variations (de Castro Tayer and Rodrigues, 2021; Wu et al., 2016, 2015); (d) delimitation of sinks based on hydrographic basins; (e) object-based image analysis (OBIA) (Dou et al., 2015); and (f) logistic regression (Kim et al., 2019).

The DEM-based sink depth is the most used, considering the difference between the DEM without depression (after filling the sinks) and the original DEM, where the areas with negative elevation mark relief depressions (Guimarães et al., 2005). However, the simple application of the difference generates many polygons, including other features unrelated to the karst depressions, such as stream channels, road drainage, or other anthropic structures, which must be filtered. In this context, different studies have included processing steps to differentiate polygons, considering: (a) digital processing of remote sensing images (Siart et al., 2009); and (b) morphometric analysis of the polygon (area, circularity, and shape indices, eccentricity, elongation rate, among others) (de

Carvalho et al., 2013; Miao et al., 2013). Several studies compared the use of this method with different Global DEMs (de Carvalho et al., 2013; Kakavas et al., 2018; Theilen-Willige, 2018), such as Shuttle Radar Topography Mission (SRTM) (Farr et al., 2007), Advanced Spaceborne Thermal Emission and Reflection Radiometer - Global DEM (ASTER-GDEM) (Tachikawa et al., 2011), Advanced Land Observing Satellite - Global Digital Surface Model (ALOS-GDSM) (Tadono et al., 2015), and Digital Surface Model using Synthetic Aperture Radar (DSM-S1) However, in areas with dense forest cover, high-density laser aerial scanning offers an alternative to overcome vegetation obstruction, allowing the underlying soil topography to be exposed. Thus, many studies have combined the DEM-based sink depth method with high-density aerial laser scanning data to obtain high-resolution topographic profiles of bare soil (Ciglič et al., 2022; Moreno-Gómez et al., 2019; Parise et al., 2020; Telbisz et al., 2016; Wall et al., 2017; Zhu et al., 2014; Zumpano et al., 2019). Nevertheless, Light Detection and Ranging (LiDAR) has the following limitations: high cost, the method not readily available, and slower due to the needs of field campaigns (Pasquetti et al., 2019).

DL-based segmentation techniques have three subdivisions: (a) semantic segmentation (class-aware labels), (b) instance segmentation (instance-aware labels), and (c) panoptic segmentation. Semantic segmentation identifies and groups pixels that are semantically together, allowing the integration of all parts of the object to form a whole, considering variations in colors, textures, shapes, and location (Garcia-Garcia et al., 2018; Guo et al., 2018). Several reviews on semantic segmentation highlight deep network architectures, data pre-processing, augmentation, algorithm taxonomy, applications, and challenges (Garcia-Garcia et al., 2018; Geng et al., 2018; Guo et al., 2018; Lateef and Ruichek, 2019; Yu et al., 2018). A limitation of semantic segmentation is the inability to distinguish different instances within the same category, not allowing to separate objects individually. Instance segmentation overcomes this limitation, discerning different labels for objects of the same category and favoring counting the number of objects and identification in conditions of occlusion or contact. Therefore, instance segmentation establishes two labels for each pixel: the object's category label and the label for each instance. Different instance segmentation models have been proposed: Fully Convolutional Instance-Aware Semantic Segmentation (FCIS) (Li et al., 2017), Mask-Region-based Convolutional Neural Network (Mask R-CNN) (He et al., 2020), Cascade Mask R-CNN (Cai and Vasconcelos, 2018; Chen et al., 2019), Mask Scoring R-CNN (Huang et al., 2019), and High-Quality Instance Segmentation Network (HQ-ISNet)

(based on Cascade Mask R-CNN) (Su et al., 2020). Reviews on instance segmentation describe the specificities and advances in this area of knowledge (Gu et al., 2022; Hafiz and Bhat, 2020; Tian et al., 2021). Using instance segmentation in satellite image processing requires some specificities due to the image size and the number of spectral bands (Carvalho et al., 2021). Finally, panoptic segmentation is a combination of semantic and instance segmentation, subdividing the target types into "things," well-defined and countable objects (such as houses, cars, and buildings), and "stuff," amorphous features of the background landscape (such as paths, lawns, and lakes). This segmentation performs in the entire scene, allowing its complete understanding (Li and Chen, 2022). In remote sensing, only some studies still use this approach (de Carvalho et al., 2022a, 2022b).

The extraordinary progress of DL using Convolutional Neural Networks (CNN) has repercussions in several studies of geomorphological pattern identification using satellite imagery, DEM, or a combination of these data: (a) aeolian landforms (Du et al., 2022; Gao et al., 2021; Shumack et al., 2020), (b) loess landforms (Li et al., 2020; Na et al., 2020), (c) landslide features (Catani, 2021), (d) permafrost landforms (Bhuiyan et al., 2020; Huang et al., 2022), (e) thermokarst landforms (Huang et al., 2018), (e) artificial landforms (terraces) (Zhao et al., 2021), (f) Martian landforms (planetary sciences) (Jiang et al., 2021; Palafox et al., 2017; Wright et al., 2022), and (g) geomorphological maps (Buscombe and Ritchie, 2018; Du et al., 2019; Li and Hsu, 2020; Meij et al., 2022; Xu et al., 2022). In mapping karst areas, few studies use deep-learning methods, and among these, all used semantic segmentation with different types of input data. These studies had specific objectives and different types of images, such as: (a) identifying early sinkholes using thermal camera images captured by drones (Lee et al., 2016); (b) detecting paleokarst collapses using 3-D seismic images (Wu et al., 2020); (c) identifying and tracking sinkholes via thermal camera video streaming (Vu et al., 2020); (d) detecting sinkholes from LiDAR-derived DEMs (Rafique et al., 2022); and (e) mapping of karst cones (positive features) from Landsat and DEM images (Fu et al., 2021). Therefore, applying deep-learning methods to sinkhole detection is still an active area of research, and to our knowledge, no study has used instance segmentation.

This research aims to evaluate the detection of karst depressions using the semantic and instance segmentation in the Central Brazil region constituted by carbonate rocks of the Bambuí Group. The secondary objectives are fourfold: (1) DEM evaluation for detection of karst depression considering SRTM, ASTER-GDEM, ALOS World 3D

(AW3D30), National Aeronautics and Space Administration DEM (NASADEM), and Copernicus 30m global DEM (GLO-30); (2) evaluation of different CNN architectures for semantic segmentation (U-Net, U-Net++, DeepLabV3+, LinkNet, and Feature Pyramid Network - FPN) and instance segmentation (Mask-CNN); (3) evaluation of box-free instance segmentation approach from semantic segmentation; (4) assessment the use of geomorphometric attributes in detection considering two approaches: (a) DEM and DEM-based sink depth; and (b) DEM, DEM-based sink depth, and nine geomorphometric attributes; and (5) elaboration of a new karst depression dataset with a significant amount of data for semantic and instance segmentations.

2. STUDY AREA

The study area is located in the region of Western Bahia (**Figure 1**), composed in its west of a plateau developed on the Urucuia group sandstones and in the east of a topographically lower terrain composed of pelitic-carbonatic rocks of the Bambuí Group. The climate is tropical, with average annual precipitations varying significantly from 1800 mm in the western part (plateau regions) to 800 mm in the eastern part (karst regions).

Considering the Brazilian speleological provinces, regions with geological formations susceptible to developing karst actions, the area belongs to the Bambuí Province (Carvalho Júnior et al., 2008; Karmann and Sánchez, 1979). This speleological province consists of Neoproterozoic sedimentary rocks of the Bambuí Group, constituted by a thick marine pelite-carbonate succession. The Bambuí Group's depositional environment was an epicontinental basin, possibly a foreland basin, positioned on the west side of the São Francisco Craton during the Brasília Belt orogeny (Coelho et al., 2008; Martins-Neto et al., 2001; Martins-Neto, 2009; Santos et al., 2000), dated at 630 Ma (Pimentel et al., 2011). Some evidence also suggests a contribution from sediments from the Araçuaí Orogeny in the Bambuí Basin. This contribution can be interpreted as restricted to the highest units of the Bambuí Group (Martins-Neto et al., 2001) or represent a longer period of involvement, as evidenced by the Três Marias Formation with sandstones zircons as young as 580 Ma (Kuchenbecker et al., 2015); and Sete Lagoas Formation with the presence of zircons of 550 Ma (Paula-Santos et al., 2015) and the Ediacaran guide fossil Cloudina sp. (Warren et al., 2014). The Bambuí stratigraphy describes three regressive cycles, each starting with a sudden, widespread influx of seawater (marine transgression) resulting from basin subsidence evidenced by deep

pelitic marine facies, followed by the shallow-platform and tidal facies and culminating in the formation of supratidal facies (Sial et al., 2009). A high density of karst depression is associated with carbonate rocks. The karst depressions form lakes that, without water, present grassland cover. The low vegetation cover has little interference in the elaboration of the DEM from different types of sensors, contrary to forest areas.

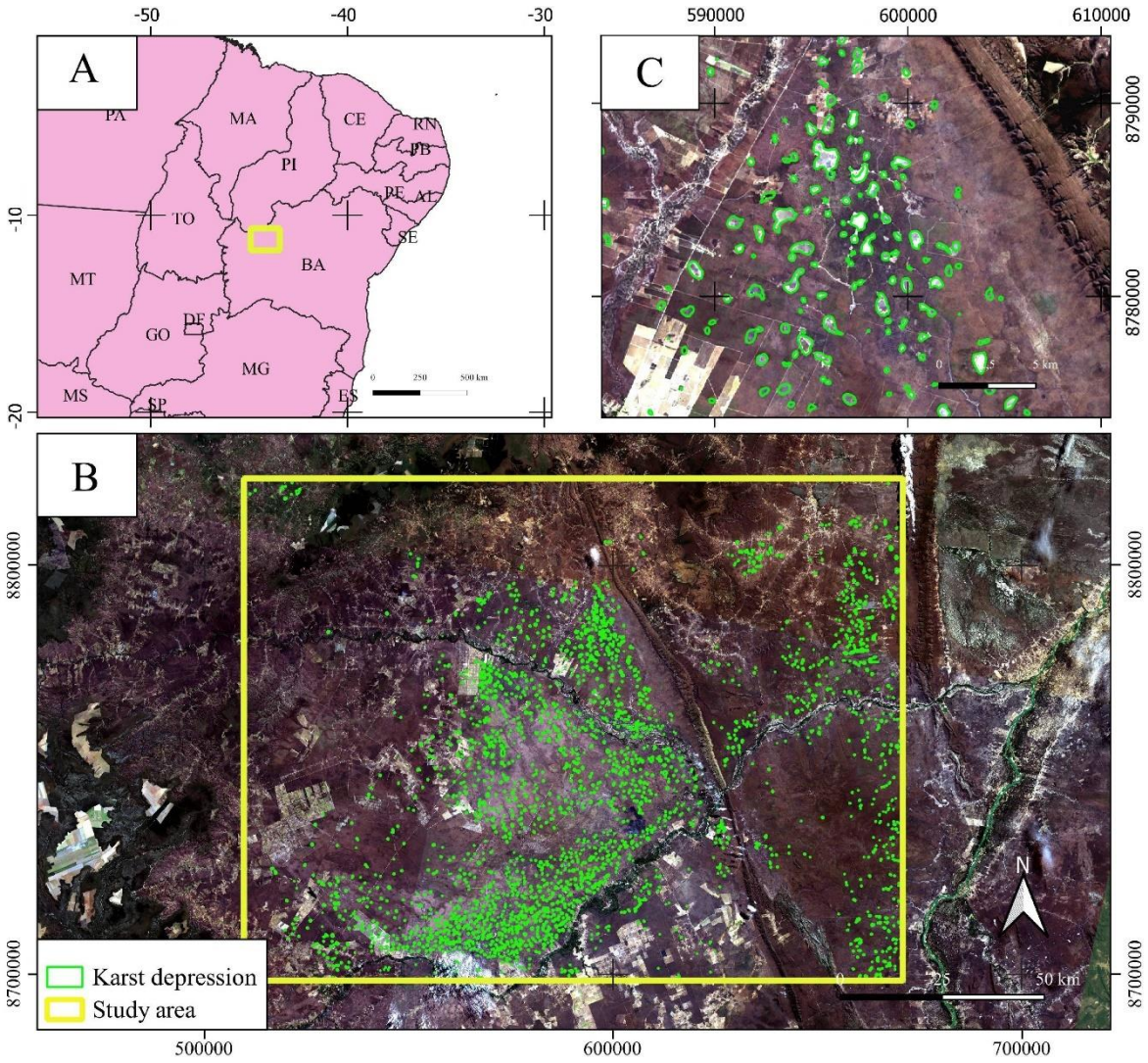

Figure 1. The study area (A) location in Brazilian territory, (B) region used in the survey, and (C) zoom of karst depressions.

This region is in the ecological transition zone between the Savanna and Caatinga ecosystems, containing Dry Forest vegetation composed of deciduous species. This vegetation has two distinct extracts: (a) the forest, with its upper canopy around 15 to 18 meters in height, regularly dense; and (b) the shrubs and young arboreal individuals, forming a dense tangle with a predominance of thorny species (Ribeiro and Walter, 2002).

Besides, the West Bahia region presents an explicit environmental and geomorphological control in the production system (Gurgel et al., 2013). The portion in the plateau areas formed by sandstones of the Urucuia Group shows an intense expansion of mechanized agriculture and landscape fragmentation (de Oliveira et al., 2017a, 2017b). In contrast, the karst regions have more preservation due to the undulation relief, soil poor in nutrients, and greater water scarcity, which inhibits agricultural occupation and favors extensive livestock.

3. MATERIAL AND METHODS

The methodology had the following steps: (a) acquisition for the study area of the following DEMs: SRTM, ASTER-GDEM, AW3D30, NASADEM, and GLO-30; (b) extraction of the DEM-based sink depth and its geomorphometric attributes (Slope, Aspect, Profile Convexity, Plan Convexity, Longitudinal Convexity, Cross-Sectional Convexity, Minimum Curvature, Maximum Curvature, and Root Mean Square - RMS) for each acquired DEM; (c) manual interpretation of karst depressions from Sentinel-2 and Operational Terra Imager (OLI)-Landsat 8 images; (d) elaboration of the Karst Depression Dataset considering training, validation, and testing samples with dimensions 128x128 considering two channels (DEM and sink depth based on DEM) and eleven channels (the two previous ones and nine morphometric attributes); (e) semantic and instance segmentation; (f) per-pixel accuracy analysis; (g) large-image reconstruction using a sliding window for semantic and instance segmentation, (h) elaboration of the instance segmentation models from semantic segmentation and GIS, and (g) per-polygon accuracy analysis. **Figure 2** presents a flowchart containing the data and the methodological steps and comparisons performed.

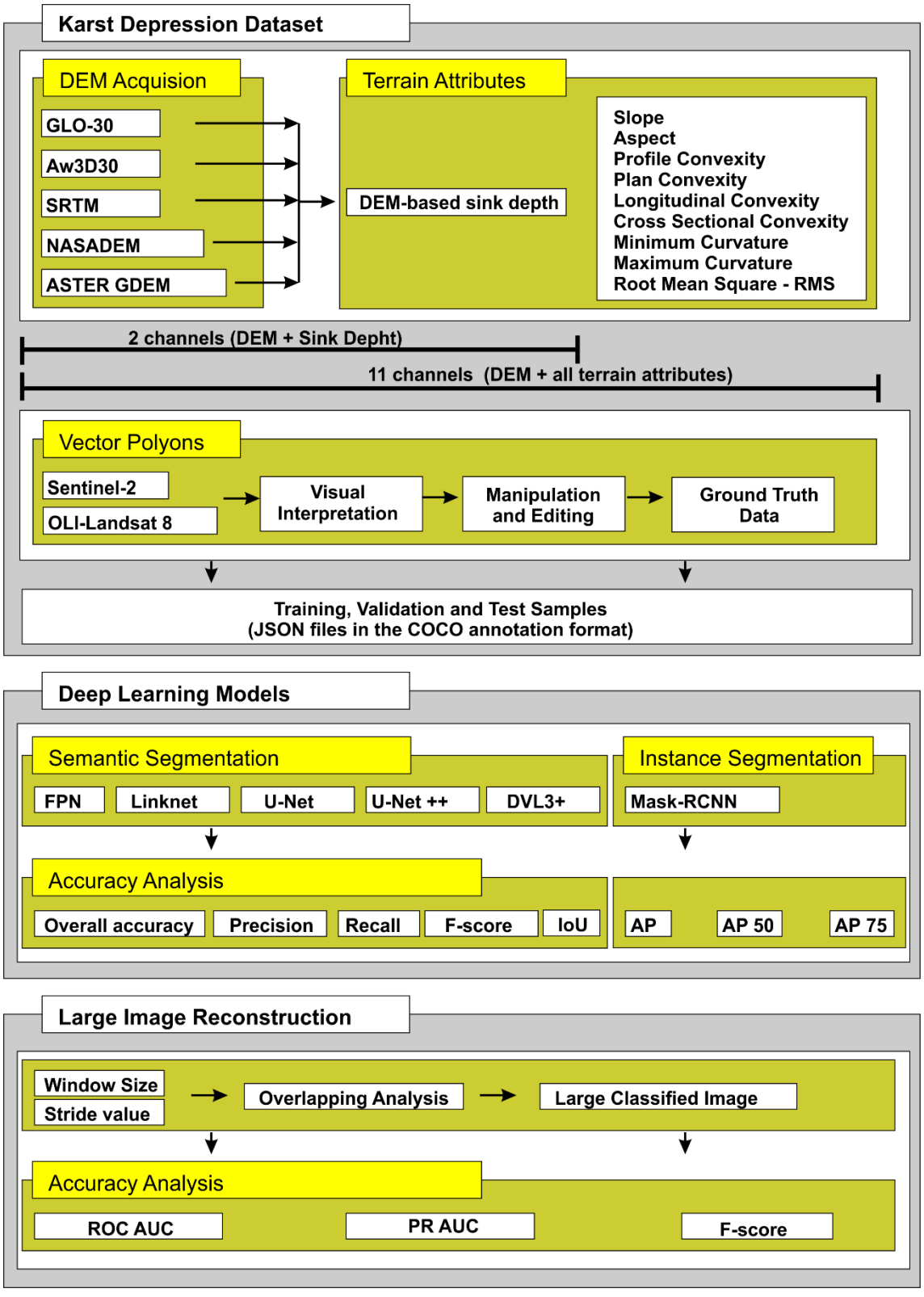

Figure 2. Methodological flowchart.

This study comprehensively compared global DEMs and semantic and instance segmentation techniques for detecting karst depressions. Semantic segmentation focuses

on classifying each pixel in an image into predetermined categories, allowing the identification of karst depressions as a distinct class. On the other hand, instance segmentation not only assigns pixels to classes but also differentiates individual instances of the same class, providing a more granular understanding of the distribution of karst depressions. By examining several state-of-the-art DL models, we investigate the performance and applicability of these segmentation methods to detect karst depressions. This comparative analysis identifies the strengths and limitations of each approach and guides the selection of the most appropriate procedure for detecting karst depressions. For both methods, the study compares different image compositions using the same field data to assess which performs better.

3.1 Digital Elevation Models and Terrain Attributes

DEMs are of fundamental importance in geomorphology for acquiring detailed information about the terrain, with applications in identifying and analyzing landforms, calculating topographic indices, identifying earth surface changes over time, and landform process modeling (Xiong et al., 2021). Different sensors provided elevation data on an almost global scale, revolutionizing the knowledge of the Earth's topography.

The SRTM mission data developed cooperatively by NASA, the National Geospatial-Intelligence Agency (NGA), the German Space Agency (DLR), and the Italian Space Agency (ASI) generated DEM from the C/X-Band Synthetic Image (SIR-C/X-SAR) with two antennas aboard the Space Shuttle Endeavor (Farr et al., 2007). Data acquisition covered the continental areas between 60 degrees North and 56 degrees South over 11 days in February 2000 (Rabus et al., 2003). The first version of the SRTM consists of the original data. The second version of the SRTM, developed by cooperation between NASA's Jet Propulsion Laboratory (JPL) and the National Geospatial-Intelligence Agency (NGA), removed artifacts (peaks and pits) and improved the definition of water bodies and coastlines; but some of the voids remain (Slater et al., 2006). The data resolution was three arc-seconds (about 90 meters) for global coverage and one arc-seconds (about 30 meters) for the United States. The third version of SRTM (SRTM v3.0) filled gaps with data from ASTER GDEM, United States Geological Survey (USGS) Global Multi-resolution Terrain Elevation Data (GMTED2010), and the USGS National Elevation Dataset (NED). Finally, a spin-off product of SRTM v3.0 was the Global 1-arc second (SRTMGL1) with high resolution (~30 m) for 80% of Earth's mass with

worldwide distribution in 2015. Complementarily, other research institutions independently propose other versions of the SRTM seeking additional improvements, such as CGIAR Consortium for Spatial Information SRTM (CGIAR-CSI SRTM v4.1) (Reuter et al., 2007), Africa Soil Information Service: Hydrologically Corrected/Adjusted SRTM DEM (AfrHySRTM) (Vagen, 2010), Elevation Corrected for Altimetry (ACE2) (Berry et al., 2010), EarthEnv-DEM90 (Robinson et al., 2014), Jonathan de Ferranti´s SRTM (Ferranti, 2021), and Multi-Error-Removed Improved-Terrain (MERIT) (Yamazaki et al., 2017).

The ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer) is a joint project sponsored by METI (Japan's Ministry of Economy, Trade, and Industry) and NASA (U.S. National Aeronautics and Space Administration). The sensor was launched onboard Terra Satellite and began to collect Earth's data in December 1999, creating data by photogrammetric processing near-infrared (NIR), wavelength (0.78–0.86 μm) with 15-m resolution (Florinsky et al., 2018; Grohmann, 2018). The ASTER GDEM is the only DEM that covers approximately 99% of Earth's surface, ranging from 83º N – 83º S, with its data generated from optical images collected between 2000-2008 and 1.2 million stereo scenes (Abrams et al., 2020; Florinsky et al., 2018; Tachikawa et al., 2011; Yue et al., 2017). Versions V1, V2, and V3 of ASTER GDEM became available in 2009, 2011, and 2019, respectively, with the latest version having a decrease in elevation void area (Abrams et al., 2020; Carrera-Hernández, 2021).

NASADEM is not a sensor launched onboard a satellite, but NASA's near-global digital elevation model derived from the reprocessing SRTM raw radar data merged with ASTER GDEM elevations, ICESat (Ice, Cloud, and Land Elevation Satellite) and GLAS (Geoscience Laser Altimeter System) data (Crippen et al., 2016). The main objective of NASADEM, released in November 2020, was to eliminate voids and other inconsistencies in the SRTM dataset, providing a more accurate product. Vertical accuracy analysis in the Mexico region demonstrated that NASADEM improved over SRTM V3, mainly in flat areas (Carrera-Hernández, 2021). However, another study in four countries (Estonia, China, New Zealand, and Norway) verified that NASADEM showed a slight improvement compared to SRTM (Uuemaa et al., 2020).

ALOS is a satellite launched by JAXA (Japan Aerospace Exploration Agency) in January 2006 and operated for five years, producing 6,5 million images covering the entire planet. The ALOS AW3D30 generation used optical stereo images from Panchromatic Remote-Sensing Instrument for Stereo Mapping (PRISM) sensor carried

by ALOS with 2,5 m resolution at nadir (NDR), forward (FWD), and backward (BWD) (Tadono et al., 2015; Takaku et al., 2018).

GLO-30 and GLO-90 data are Copernicus DEMs (COP-DEMs) with 30-m and 90-m resolutions derived from TerraSAR-X add-on for Digital Elevation Measurement (TanDEM-X) data collected between 2011 and 2015 (ESA, 2023) and freely available on the ESA website. These data have been considered accurate cutting-edge global DEMs (Hawker et al., 2022). According to Guth and Geoffroy (2021), the GLO-30 demonstrates superiority compared to other global 1″ (second of arc) DEMs (ALOS, ASTER, NASA, and SRTM) evaluated with LiDAR and ICESat-2 data. In geomorphology studies, Copernicus DEM detected halo-karst development on the shoreline of the Dead Sea (Closson et al., 2023) and flood-prone areas (Cuellar et al., 2022).

Besides the DEM, this research evaluated the insertion of terrain attributes as additional channels to highlight features and better segmentation of karst depressions. Therefore, the analysis included the DEM-based sink depth generated by each model due to its importance and wide use in previous studies of karst depression detection (Antonić et al., 2001; Guimarães et al., 2005). Complementarily, the study also extracted the geomorphometric attributes from the DEM-based sink depth for the enhancement of karst depressions: (a) slope; (b) aspect; (c) profile convexity; (d) plane of convexity; (e) longitudinal convexity, (f) cross-sectional convexity, (g) minimum curvature, (h) maximum curvature, and (i) root mean square (RMS) (Wood, 1996). **Figure 3** exemplifies the geomorphological attributes used in detecting karst depressions for a small part of the study area. This example uses the GLO-30 image's terrain attributes (**Figures 3A to 3K**). **Figure 3L** represents the visual interpretation of the karst depressions used as field truth data in the deep segmentation processing.

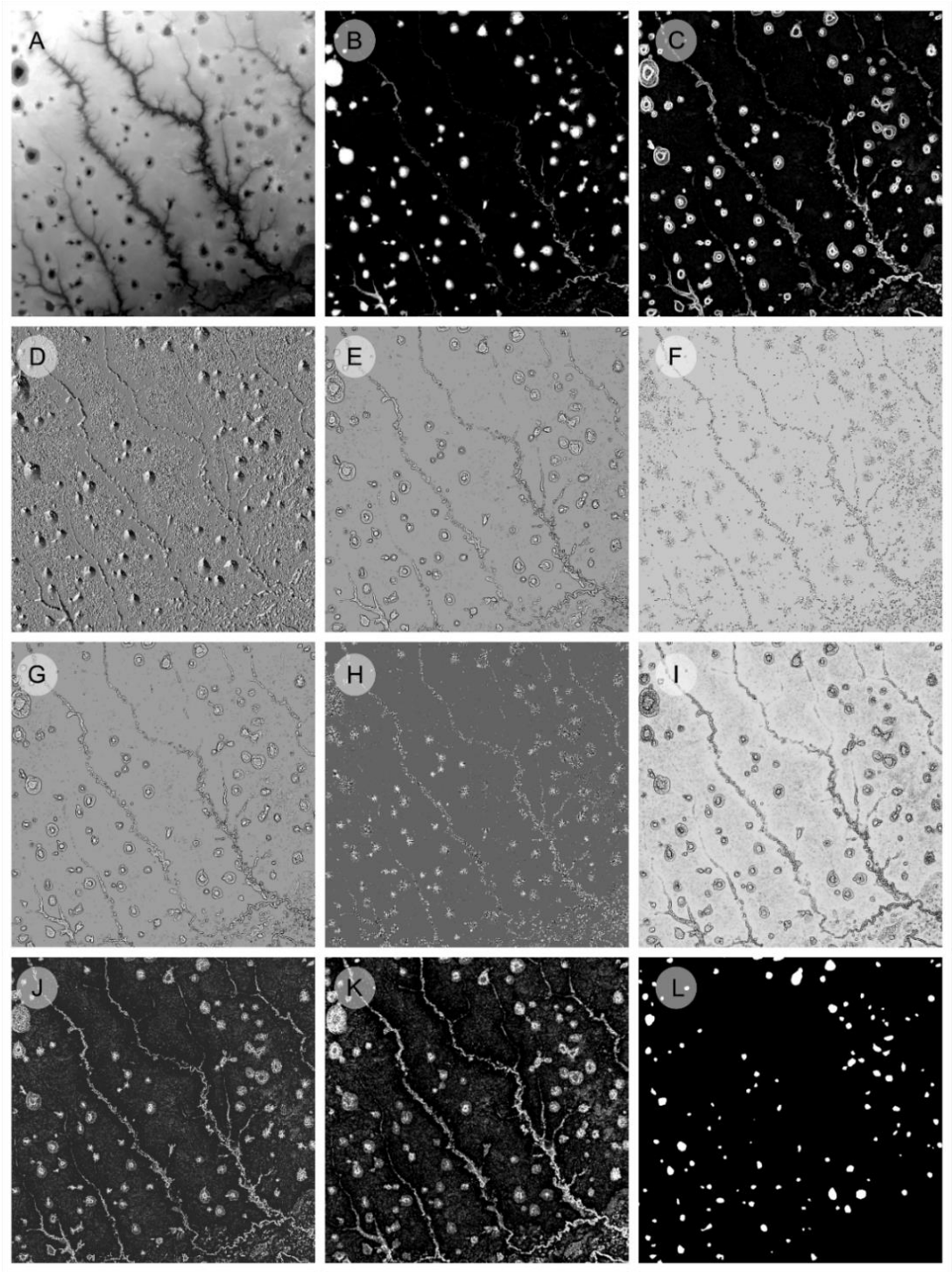

Figure 3. Exemplification of terrain attributes used in the detection of karst depressions at a zoom of 800 x 800 pixels contained in the study área: : (A) GLO-30, (B) DEM-based sink depth, (C) slope; (D) aspect; (E) profile convexity; (F) plane of convexity; (G) longitudinal convexity, (H) cross-sectional convexity, (I) minimum curvature, (J) maximum curvature, (K) root mean square (RMS), and (L) ground truth based on visual interpretation. In the present example, the elaboration of all-terrain attributes used the GLO-30 image.

Therefore, the study compared two data sets for each of the evaluated DEMs: (a) DEM and DEM-based sink depth (2 channels), and (b) DEM, DEM-based sink depth, and nine terrain attributes (11 channels). An essential issue is that previous tests using only the DEM (without including terrain attributes) presented very discrepant errors not being considered in the analysis.

3.2. Karst Depression Dataset

There is a close relationship between the advancement of DL tasks and the provision of large datasets for testing. However, few geomorphological datasets are available for DL processing due to its recent introduction in this field of science. Therefore, elaborating geomorphological datasets is crucial in encouraging other researchers to develop new methods with greater precision and drive the progress of geomorphological mapping. In this context, the present study developed a database of karst depressions to meet both semantic and instance segmentation.

Establishing criteria for delimiting karst depressions is a significant challenge to creating a good dataset. In the present study, field truth annotation of karst depressions used visual interpretation from the Sentinel-2 and OLI-Landsat 8 images. Therefore, karst depression annotations present additional difficulties for visual interpretation, considering the following factors: (a) amorphous shapes; (b) size variation; (c) difficulty in demarcating the exact limit in places with tenuous changes; (d) the limit may show variations; and (e) similarities with other geomorphological features such as river features. The on-screen visual interpretation of the images used the Quantum Geographic Information System (QGIS) program, which has numerous tools for drawing, manipulating, and editing the polygons referring to the karst depressions.

The elaboration of the training, validation, and test samples considered the ground truth image obtained by visual interpretation and morphometric attributes composed of 2 and 11 channels for each of the five DEMs evaluated. Consequently, we developed ten types of sample sets considering combinations of five DEMs and two types of morphometric attribute sets. The preparation of the samples used the module developed within the free *Abilius* software that converts the input data to JSON files in the COCO annotation format with wide use in DL (https://github.com/abilius-app/Panoptic-Generator) (de Carvalho et al., 2022b). In an end-to-end approach, this tool can generate annotations for the three segmentation types (semantic, instance, and panoptic) and

consider multichannel without restricting the input data to the three bands (Reed, Green, and Blue - RGB). The tiles' spatial dimension was 128 x 128 pixels, cropped in a quadrangular window positioned from central vector points chosen in the study area, considering depression features, flat areas, and similar features such as oxbow lakes that generate depressions as its water evaporates. Most karst depressions have isolated circular or elliptical shapes. The dataset division to evaluate the performance of segmentation models considered 1600 samples for training, 400 samples for validation, and 400 samples for testing.

3.3. Semantic Segmentation Models

The semantic segmentation analysis compared five architectures: U-net (Ronneberger et al., 2015), U-net++ (Zhou et al., 2018), DeepLabV3+ (Ronneberger et al., 2015), LinkNet (Chaurasia and Culurciello, 2017), and Feature Pyramid Network (FPN) (Lin et al., 2017). The U-net architecture is a convolutional network initially designed for the segmentation of biomedical images and which currently has a wide diffusion. This architecture features a symmetric encoder-decoder structure that captures high-level semantic information and low-level spatial details (Ronneberger et al., 2015). U-net++ is a modified version of the original architecture that introduces nested skip pathways and dense connectivity to improve gradient flow and enhance feature reuse (Zhou et al., 2018). The DeepLabV3+ architecture combines an atrous spatial pyramid pooling (ASPP) module with an encoder-decoder structure to capture multi-scale contextual information and achieve precise segmentation (Chen et al., 2018). The LinkNet architecture is a lightweight, real-time semantic segmentation network that employs skip connections between encoder and decoder layers to facilitate feature fusion and maintain spatial resolution (Chaurasia and Culurciello, 2017). Lastly, the FPN architecture employs a top-down pathway to construct high-level semantic feature maps at various scales, enhancing multi-scale feature learning and prediction. The backbone used in all architectures was Efficient-net-b7 (Tan and Le, 2019), which balances computational efficiency and model capacity, resulting in greater accuracy.

To maintain consistency among the semantic segmentation models, we used the same hyperparameters during the training procedure: (1) a learning rate of 0.001, (2) the number of epochs of 150, (3) a batch size of 25, (4) Adam optimizer, (5) sigmoid activation function in the last layer. Besides, the training of all models used 128x128-

sized images, and the number of channels depended on the images, which could be 2 or 11. To avoid overfitting, we saved the model with the lowest binary cross-entropy loss within the validation set.

3.4. Instance Segmentation Models

The instance segmentation analysis utilized the Mask-RCNN architecture that combines region proposal networks (RPN) and Fast-RCNN to differentiate and segment individual objects within an image (He et al., 2020). For this stage, we incorporated the ResNeXt-101 backbone, an advanced variant of the ResNet architecture that leverages cardinality and split-transform-merge operations to enhance representational capacity and promote training efficiency. The instance segmentation models present different parameters in the architecture, which affects the training procedure. Mask-RCNN model training considered the following configuration: 15,000 iterations (with a thousand warm-up iterations), a learning rate of 0.0001, gamma of 0.1, 500 ROIs per image, anchor sizes of [8,16,32,64,128], and one image per batch. The evaluation of the validation set occurred after 3,000 iterations, considering the saved model had the lowest total loss.

3.5. Image Reconstruction using Sliding Window Approach

Semantic segmentation models generate predictions for a specific patch size, typically with 2^n square dimensions (e.g., 16, 32, 64, and 128), in our case, 128x128 pixels. Remote sensing orbital satellite images are usually huge, and a sliding window approach is necessary to classify more significant regions effectively. The concept is to classify smaller frames, one at a time, and move the window along the x and y-axis. This window can move at a smaller or larger rate, defined by the stride value, which is the number of pixels the window moves from one prediction to another. When the stride value is smaller than the window size, there is an overlap in the predictions, and a pixel assumes different values. Therefore, establishing the threshold value of positive predictions to perform the segmentation is another critical analysis.

Previous studies reported that the smaller the stride, the better the metrics, mainly due to the improved classification of the center of each frame than the boulders, creating discontinuities attenuated by averaging overlapping pixels (de Albuquerque et al., 2020). Typically, targets sectioned at the edges of frames have higher errors due to not containing

the entire feature of interest, which are corrected by shifting the sliding window to the center in subsequent frames. However, a higher number of frames increases the computational demands, requiring adjustments that align the sliding window stride value with the available computational resources and the specific requirements of the problem.

Many studies establish, as the number of positive predictive values in the window sequence, a threshold value corresponding to the mean of the accumulated data. The present study evaluates, from the best training model, the following parameters for the reconstruction of large images: (a) stride value for the moving window comparing 2^n sizes (4, 8, 16, 32, and 128) and (b) threshold value for the resulting probabilistic image of the different stride values.

The methodology for the instance segmentation model is different from semantic segmentation since the predictions have unique values. The overlapping pixels would have different values, meaning the average would be ineffective. Besides, the instance segmentation models use bounding boxes, which introduces another complexity. We used the mosaicking approach to solve this problem developed by Carvalho et al. (2021). This methodology consists of classifying objects in the border of each frame using an intermediate window. Then, we apply a non-maximum suppression algorithm to maintain the larger objects representing the complete predictions, eliminating partial predictions for the same object.

3.6. Elaboration of Semantic-to-Instance Segmentation Conversion using GIS

In addition to employing instance segmentation methods using the Mask-RCNN, the present research used another approach to obtain box-free instance-level information from semantic segmentation data and the Geographic Information System (GIS). Therefore, semantic-to-instance segmentation conversion is a computer vision task that involves transforming a semantic segmentation map into an instance segmentation map. This approach is an evolving research area with different possible algorithms depending on the specific requirements and constraints of the problem. This conversion can be done considering two different situations: (a) targets in contacts, which explored an edge to isolate the interior of objects using multitasking learning (Mou and Zhu, 2018) or multiclass learning (de Carvalho et al., 2022c) procedures to avoid agglutination, and (b) targets that are not in contact, where the polygonization of semantic features creates a specific identifier per instance (de Carvalho et al., 2023). The karst depressions insert on

the approach where there is no interaction of the features, facilitating the conversion of semantic resources into instances. This free bounding box instance segmentation approach has some advantages over the traditional instance segmentation method (de Carvalho et al., 2023): (a) it does not require data with such structured information; (b) fewer parameters for detecting objects (i.e., anchor boxes) and procedures (i.e., controlled ROI); (c) the image reconstruction process by sliding windows is more straightforward, for example, using only the average value between the overlapping windows; and (d) better per-pixel accuracy metrics, especially for small objects.

3.7 Accuracy Metrics

The accuracy analysis adopted per-pixel accuracy metrics obtained by the confusion matrix. The primary per-pixel metrics for semantic segmentation are Overall Accuracy, Precision, Recall, F-Score, and Intersection over Union (IoU) (**Table 1**). Overall Accuracy is the ratio of correctly classified pixels to the total number of pixels. Precision is the fraction of correctly predicted positive instances over the total number of predicted positive instances. Recall, also known as sensitivity, is the ratio of correctly predicted positive instances over the total number of true positive instances. F-score is the harmonic mean of precision and recall, providing a balanced performance measure. Intersection over Union (IoU) calculates the overlap between the predicted and ground truth segmentation, consisting of the ratio of the intersection area to the union area of the predicted and ground truth segmentations. These metrics enable us to assess the segmentation performance comprehensively, allowing for a thorough analysis of the proposed method's effectiveness in various scenarios.

The research also evaluates the instance segmentation performance using the COCO metrics based on Average Precision (AP) and its variants, AP50 and AP75. These metrics are widely adopted in the computer vision community to assess the quality of instance segmentation models. AP is the mean precision value calculated at different Intersection over Union (IoU) thresholds (ranging from 0.5 to 0.95 with a step of 0.05). AP offers an overall assessment of the model's performance by considering both localization accuracy and classification performance across a range of IoU thresholds (**Table 1**). AP50 and AP75 are two specialized variants of the AP metric, calculated at fixed IoU thresholds of 0.5 and 0.75, respectively. These metrics provide insights into the

model's performance at different levels of segmentation precision, ranging from AP50 to AP75, from least to most stringent.

Table 1. Accuracy metrics equations, where "TP" is true positive, "TN" is true negative, "FP" is false positive, "FN" is false negative, "IoU" is Intersection over Union, "N" is the number of IoU thresholds, and "t" is a specific IoU threshold.

Metric	Equation
Overall accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$
Precision	$\frac{TP}{TP + FP}$
Recall	$\frac{TP}{TP + FN}$
F-score	$\frac{2\ x\ (Precision\ x\ Recall)}{Precision + Recall}$
IoU	$\frac{TP}{TP + FP + FN}$
AP	$\frac{1}{N}\sum(Precision@IoU\_t)$

Complementarily, we evaluated the reconstructed images of the semantic and instance segmentations from the Receiver Operating Characteristic (ROC) and Precision-Recall (PR) curves and their respective Area Under Curve (AUC) scores. ROC curve summarizes the binary model's performance, produced by the graph between the Rate of False Positives (or Sensitivity, Recall) arranged along the x-axis versus the Rate of True Positives (or Specificity) on the y-axis, described by TN/ (TN+ FP). Therefore, considering different classification threshold values, the ROC curve provides information about the probabilistic model output of the sliding window method. The ideal performance scenario is when the Sensitivity is 1 (at the top of the graph), and the Specificity is 0 (at the left).

The Precision-Recall (PR) curve is also a graph used to evaluate classification models' performance, mainly regarding class imbalance, such as the current study where there are significantly more extensive areas without karst depressions than with karst depression. Given this pronounced class imbalance in the binary classification of dolines using probability images generated through moving windows, PR curves are better suited for analysis than ROC curves. The PR curve plots Precision (on the y-axis) versus Recall (on the x-axis), focusing on the model's ability to identify the positive class accurately. The area under the curve (AUC) in both plots provides a single value of all threshold

values, making comparing two or more doline probability images easy. The F-score determined the optimal threshold value for karst depression detection for each sliding window stride value.

4. RESULTS

4.1. Semantic Segmentation Results

This study compared the performance of semantic segmentation considering the following factors: (a) five architectures with the EfficientNet-B7 backbone, (b) five different DEMs, and (c) two sets of attributes with two channels (DEM and DEM-based sink depth) and eleven channels (DEM, DEM-based sink, and nine geomorphometric metrics). **Figure 4** exemplifies the different semantic segmentation procedures for a frame with 128x128 pixels considering the different DEMs and architectures. The different tests used the same ground truth mask with sample collection at the exact location, where the model's configurations were consistent throughout all the experiments.

Table 2 comprehensively summarizes the results for each image composition, different DEMs, and architectures. The highest precision value within each precision metric for each semantic segmentation architecture is bold, italicized, and underlined in **Table 2**, while the lowest value is bold only. The black squares are the highest values of an accuracy metric within the entire evaluated set, and the gray squares are the lowest.

Comparing the semantic segmentation architectures, the FPN showed a clear predominance of the best accuracy metrics within the same DEM and set of variables (**Table 2 – Best Models**). The FPN architecture obtained the best F-Score and IoU metrics considering each DEM evaluated with 2 or 11 channels. In addition, FPN had the best Overall Accuracy metrics, except for ASTER GDEM (2 variables), which presented as the best result the DLV3+. Within all the tests performed, the FPN with the GLO-30 (11 variables) obtained the highest Accuracy Overall, F-Score, and IoU, and with the Aw3D30 (2 variables) the highest Precision metric, losing only in the Recall metric for LinkNet with GLO-30 (11 variables). Among the tests performed, the DLV3+ architecture had the highest number of worst results for Overall Accuracy, F-Score, and IoU.

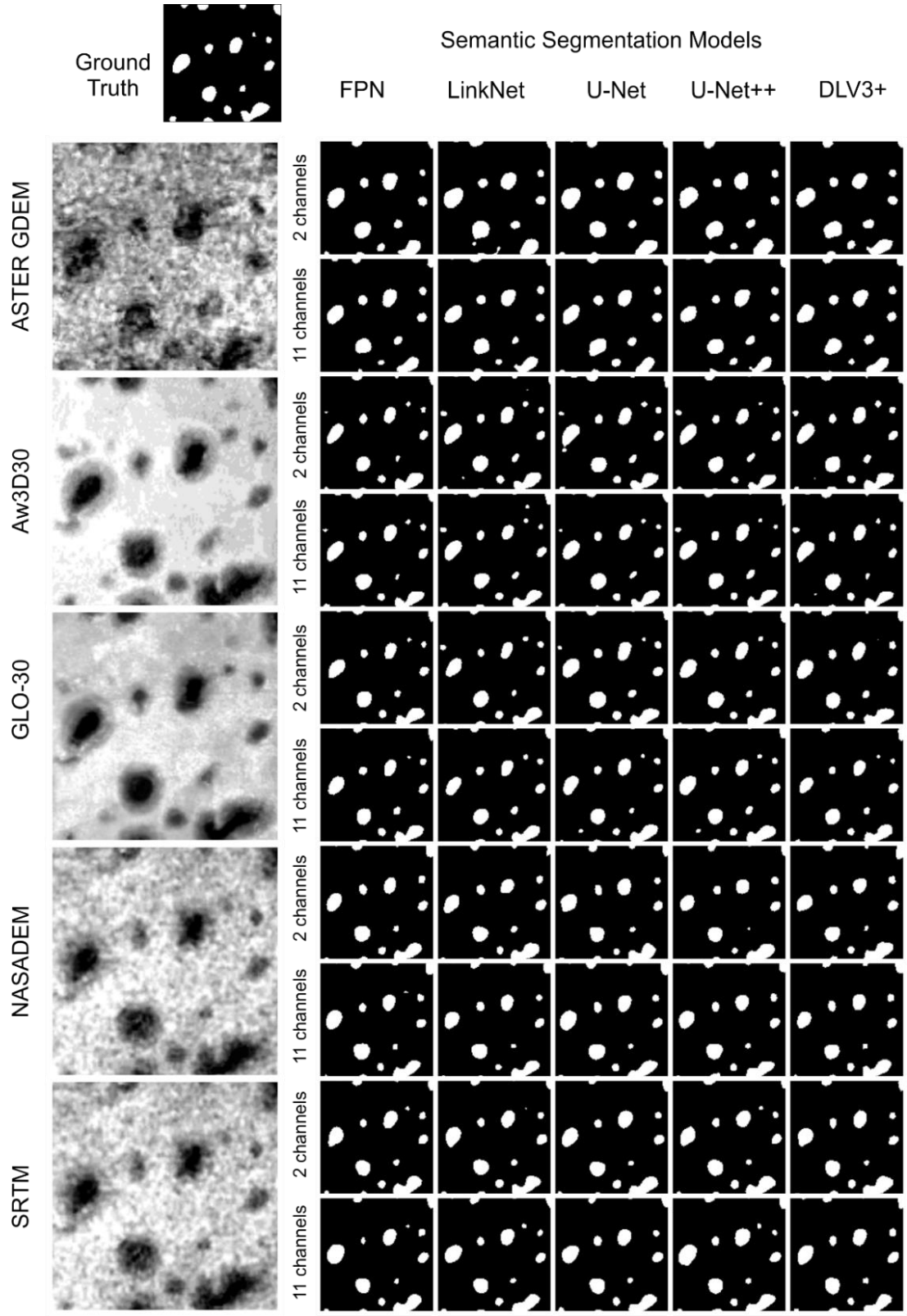

Figure 4. Exemplification of karst depression prediction using five different semantic segmentation methods and five digital terrain models.

Table 2. Semantic segmentation results considering the Overall Accuracy (OA), Precision (P), Recall (R), F-Score (FS), and Intersection Over Union (IoU) metrics for 10 distinct image compositions: (1) ASTER GDEM (2 variables), (2) ASTER GDEM (11 variables), (3) Aw3D30 (2 variables), (4) Aw3D30 (11 variables), (5) GLO-30 (2 variables), (6) GLO-30 (11 variables), (7) NASADEM (2 variables), (8) NASADEM (11 variables), (9) SRTM (2 variables), and (10) SRTM (11 variables).

	U-Net					FPN				
	OA	P	R	F-S	IoU	OA	P	R	F-S	IoU
1	97.36	71.30	**61.48**	**66.03**	**49.28**	**97.47**	74.05	**60.65**	**66.69**	**50.02**
2	**97.33**	**69.72**	63.53	66.48	49.79	**97.47**	**73.19**	62.22	67.26	50.67
3	98.40	81.47	79.95	80.70	67.65	98.52	***87.27***	75.53	80.97	68.03
4	98.52	82.84	81.61	82.22	69.81	98.57	83.68	81.73	82.69	70.49
5	98.41	80.12	***82.35***	81.22	68.38	98.53	85.73	77.85	81.6	68.92
6	***98.60***	***84.22***	81.96	***83.07***	***71.05***	***98.65***	84.83	***82.34***	***83.57***	***71.77***
7	97.73	72.34	74.14	73.23	57.76	98.13	79.2	75.08	77.09	62.72
8	98.20	79.07	77.47	78.26	64.29	98.27	82.0	75.11	78.41	64.48
9	98.31	81.68	76.71	79.12	65.45	98.39	81.49	79.28	80.37	67.18
10	98.36	80.90	79.36	80.12	66.83	98.38	81.93	78.48	80.17	66.91
	DLV3+					U-Net++				
	OA	P	R	F-S	IoU	OA	P	R	F-S	IoU
1	97.48	74.90	**59.53**	66.34	49.63	**97.44**	**73.97**	59.77	**66.12**	**49.38**
2	**97.45**	**74.09**	60.03	**66.32**	**49.62**	97.51	75.54	**59.51**	66.58	49.90
3	98.30	78.94	***80.90***	79.91	66.54	98.44	***86.32***	74.37	79.9	66.53
4	98.46	83.56	78.57	80.99	68.05	98.54	83.46	81.07	82.25	69.85
5	98.25	78.00	80.81	79.38	65.81	98.39	80.64	81.02	80.83	67.83
6	***98.61***	***85.49***	80.32	***82.83***	***70.69***	***98.58***	83.24	***82.65***	***82.94***	***70.85***
7	97.88	74.45	74.95	74.7	59.62	97.78	77.68	65.84	71.27	55.37
8	98.12	78.36	75.99	77.16	62.81	98.16	77.45	79.1	78.27	64.3
9	98.37	84.73	74.15	79.09	65.41	98.32	82.14	76.23	79.08	65.39
10	98.26	78.82	79.50	79.16	65.51	98.33	80.31	79.5	79.9	66.53
	Linknet					Best models				
	OA	P	R	F-S	IoU	OA	P	R	F-S	IoU
1	**97.26**	**69.82**	**60.60**	**64.89**	**48.02**	**DLV3+**	**DLV3+**	**U-Net**	**FPN**	**FPN**
2	97.35	70.94	62.07	66.21	49.49	**FPN**	**U-Net++**	**U-Net**	**FPN**	**FPN**
3	98.38	81.97	78.63	80.27	67.04	**FPN**	***FPN***	**DLV3+**	**FPN**	**FPN**
4	98.51	83.21	80.66	81.92	69.37	**FPN**	**FPN**	**FPN**	**FPN**	**FPN**
5	98.45	82.67	79.48	81.04	68.13	**FPN**	**FPN**	**U-Net**	**FPN**	**FPN**
6	***98.51***	80.81	***84.44***	***82.59***	***70.34***	***FPN***	**DLV3+**	***Linknet***	***FPN***	***FPN***
7	97.80	74.90	71.21	73.01	57.49	**FPN**	**FPN**	**FPN**	**FPN**	**FPN**
8	98.20	80.40	75.24	77.73	63.58	**FPN**	**FPN**	**U-Net**	**FPN**	**FPN**
9	98.36	82.51	77.02	79.67	66.22	**FPN**	**DLV3+**	**FPN**	**FPN**	**FPN**
10	98.38	**83.66**	75.96	79.62	66.15	**FPN**	**Linknet**	**DLV3+**	**FPN**	**FPN**

Comparing the DEMs, the GLO-30 achieved supremacy of the best accuracy metrics in the five tested architectures. Therefore, the best DEM for OA, F1, and IoU metrics in all architectures was GLO-30 with 11 variables, followed by Aw3D30 with 11 variables. In the third and fourth positions appear these DEMs containing two variables. Although there are variations in the Precision and Recall metrics, GLO-30 and Aw3D30 with 2 and 11 variables predominate in the first four best models in different architectures. SRTM and NASADEM models with 11 and 2 variables appear with intermediate precision results. Finally, ASTER GDEM consistently presents the worst accuracy values for all tested metrics and architectures.

By analyzing the attributes used in karst depression segmentation, the models that used only the DEM attribute obtained gross errors and were disregarded in the analysis. The combination of DEM, DEM-based sink, and terrain attributes showed higher accuracy than the simple combination of DEM and DEM-based sink (**Table 3**). Among the 125 comparisons tested in the semantic segmentation between two and eleven variables (considering the five-accuracy metrics, five deep-learning architectures, and the six DEMs), only 24 comparisons showed superior values of the models of two variables to those of eleven variables. Therefore, 80% of the metrics using eleven variables were more accurate than those with two. Precision and Recall were the metrics that concentrated superiority with two variables.

Table 3. Comparison of semantic segmentation accuracy metrics of models with two or eleven variables. Areas marked in grey boxes represent greater accuracy of semantic segmentation with 11 variables over those with 2 variables, and the white areas are the opposite. The five digital terrain models used are: (1) ASTER GDEM, (2) Aw3D30, (3) GLO-30, (4) NASADEM, and (5) SRTM.

	U-Net					FPN					LinkNet					DVL3+					U-Net++				
	OA	P	R	F-S	IoU	OA	P	R	F-S	IoU	OA	P	R	F-S	IoU	OA	P	R	F-S	IoU	OA	P	R	F-S	IoU
1																									
2																									
3																									
4																									
5																									

4.2. Results of the Semantic Segmentation Image Reconstruction

The construction of the semantic segmentation models used training samples of 128x128 pixels, with application to large remote sensing images through sliding windows. Image reconstruction considered only the FPN model using the GLO-30 DEM with 11 variables, which obtained the best results compared to the other tested models. The analysis of the reconstructed image involved the comparison of various stride values for the sliding window (4, 8, 16, 32, 64, and 128). Each stride value resulted in a probability image, where pixel values ranged from 0 to 1, with 1 indicating the highest likelihood of the target object's presence. The ROC AUC and PR AUC values allow a performance comparison between reconstructed images with different stride values.

Table 4 lists the AUC ROC and AUC PR values of the images resulting from the use of the sliding window method with the FPN model and the DEM GLO-30 with 11 variables, considering the 2^n stride values from 4-pixel stride to 128-pixel stride. The AUC values gradually improved with decreased stride values from 128 to 8. Therefore, smaller strides favor the sliding window to obtain better coverage of the object and less risk of losing information, resulting in better predictions. Larger or complexly shaped objects can be better captured with smaller increment windows, ensuring that objects are not only partially captured. Thus, the sliding window ensures obtaining the context around each pixel, incorporating more contextual cues into the segmentation process, and eliminating ambiguities. In contrast, when the stride is too large, the pixels at the edges of the windows may not contribute to the final segmentation or may only partially influence the predictions. Regarding the threshold values, the best results in the different strides were close to 0.4. At the optimal threshold point, F-score values between 4 and 32-pixel stride are equivalent, obtaining a value of 0.76.

Table 4. Accuracy metrics for semantic segmentation image reconstruction models, considering sliding window method with six lengths of strides. The accuracy metrics used were Receiver Operating Characteristic (ROC) Area Under the Curve (AUC) and Precision-Recall (PR) AUC. The highest values are italics and bold numbers.

Stride value	ROC AUC	PR AUC	Optimal threshold value	F-score
128-pixel stride	92.13	68.92	0.17	0.73
64-pixel stride	95.13	76.39	0.50	0.75
32-pixel stride	96.58	79.4	0.39	***0.76***
16-pixel stride	97.23	80.58	0.40	***0.76***
8-pixel stride	97.81	81.52	0.40	***0.76***
4-pixel stride	***98.18***	***82.32***	0.41	***0.76***

Figure 5 exemplifies image reconstruction results using sliding windows with different strides and the threshold value of 0.5 considering good and bad fit scenarios. **Figure 5A** demonstrates in the 128-pixel stride frame (no overlapping) a polygon sectioned in the middle referring to the edge effect, which does not correspond to the truth ground image. Because of the occlusion, elements on the window edges have continuity errors. Using smaller stride values gradually corrects, merging the two segments into one. **Figure 5B** demonstrates the window border effect with a 128-pixel stride vertically and

horizontally, which does not provide a good fit. Assuming smaller strides eliminates the small isolated fragments of the 128 and 64-pixel strides but erroneously generates two polygons. Furthermore, **Figure 5** demonstrates a polygon elimination in the different predictions.

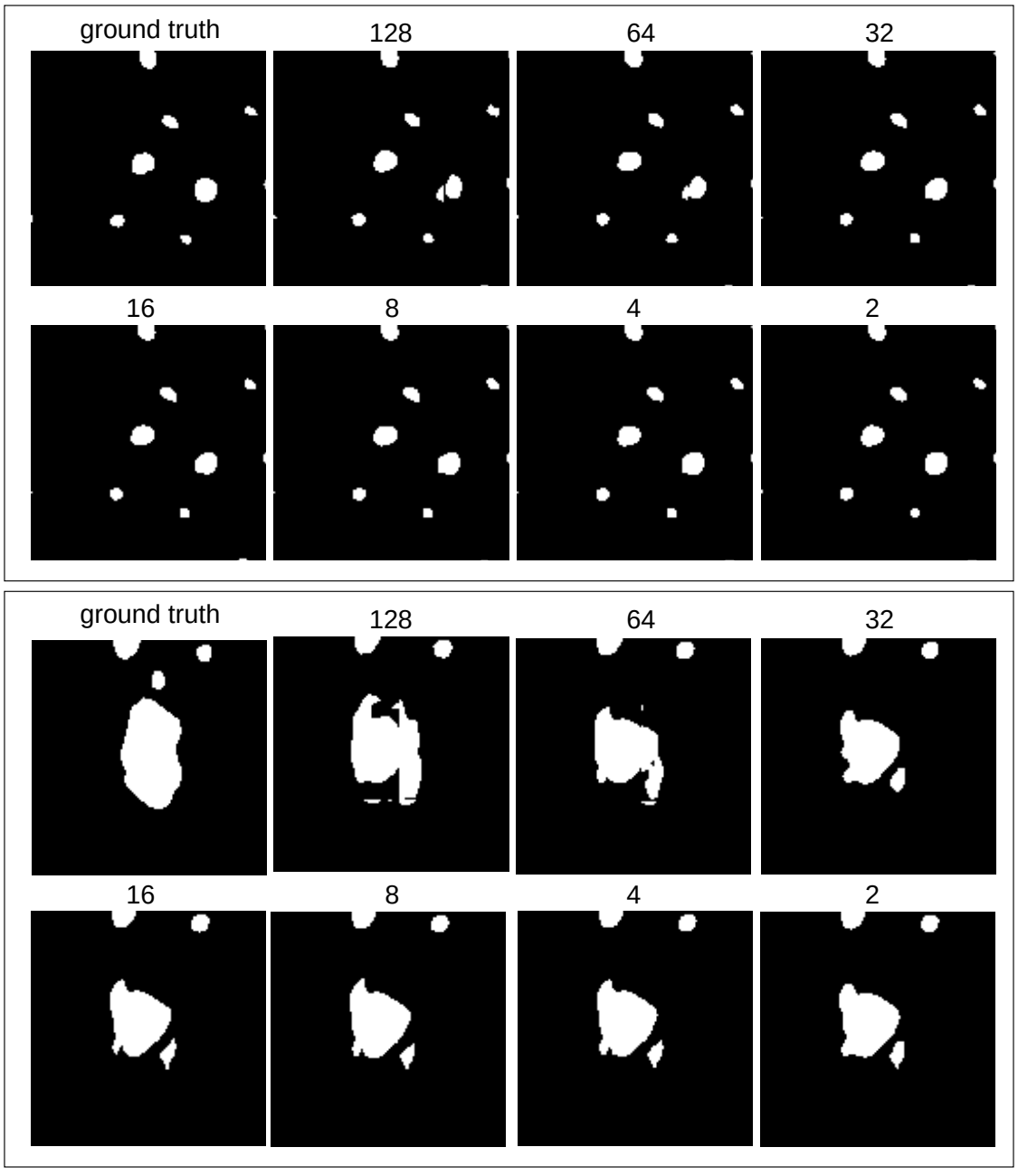

Figure 5. Exemplification of image reconstruction results using the sliding window with different stride values over areas with edge effect, considering two situations with good (a) and reasonable (b) prediction. The model considered the FPN semantic segmentation method, GLO-30, and the threshold value of 50%.

4.3. Instance Segmentation Results

The instance segmentation results using the Mask-RCNN architecture with the ResNeXt-101 backbone agree with the semantic segmentation results (**Table 5**). The GLO-30 DEM with 11 variables obtained the best AP50, reaching 83.85%, while the ASTER GDEM had the worst result.

Table 3. Common Objects in Context (COCO) metrics for the 11 image compositions, considering the Average Precision (AP), AP50, and AP75 for bounding box and segmentation mask for 10 distinct image compositions.

	Box			Mask		
Image Composition	AP	AP50	AP75	AP	AP50	AP75
ASTER GDEM (2 variables)	15.804	46.320	5.518	13.571	44.356	3.010
ASTER GDEM (11 variables)	16.327	46.458	7.306	14.812	44.160	5.808
Aw3D30 (2 variables)	30.726	78.440	13.773	31.781	79.096	16.505
Aw3D30 (11 variables)	35.679	80.854	22.727	31.616	77.985	16.181
GLO-30 (2 variables)	33.219	82.539	17.666	36.265	**83.451**	**25.555**
GLO-30 (11 variables)	**37.947**	**83.356**	**26.716**	**36.557**	81.517	24.186
SRTM (2 variables),	27.625	74.349	11.635	24.908	71.692	8.768
SRTM (11 variables).	27.284	73.567	12.242	25.879	70.983	10.817

Even though the AP metric is very relevant for competitions, for practical applications is much more straightforward to use static IoU thresholds. Since the karst depressions are very difficult to delineate the borders, stricter metrics may not be as suitable for this target, and less strict metrics can be more interesting. Another issue with using 30-m resolution DEM is that all karst depressions acquire the categorization of small objects (an area smaller than 32x32 pixels) according to Common Objects in Context (COCO) (Lin et al., 2014). Despite the importance of the subject, the detection of small objects has a little amount of research. (Tong et al., 2020).

4.4. Results of Instance Segmentation Image Reconstruction

The technique for reconstructing large images using instance segmentation employs a sliding window mechanism distinct from semantic segmentation models that rely on overlapping pixels. In the case of Mask-RCNN, which provides outputs such as bounding boxes, classification, and segmentation masks, the sliding windows implementation incorporates a non-maximum suppression step sorted by area to remove redundant detections and merge overlapping instances (Carvalho et al., 2021). The classification threshold is the primary parameter for tuning the information in the sliding

windows in this process. Adjusting the classification threshold has implications for the performance of the instance segmentation. Employing a stricter threshold increases the likelihood of false negatives, where objects may be missed or improperly classified. Conversely, adopting a less strict threshold tends to increase the occurrence of false positives, where objects are wrongly identified or segmented. It is crucial to balance these two types of errors to achieve the most accurate instance segmentation results. Therefore, we compared five thresholds considering accuracy metrics: 70%, 75%, 80%, 85%, and 90% (**Table 6**). In our evaluation, the threshold value of 90% emerged as the best compromise between false negatives and false positives, achieving the highest value of F-Score. **Figure 6** demonstrates the result of the instance segmentation containing the masks and bounding boxes of the karst depressions. The segmentation result ignores river features (such as abandoned meanders, oxbow lakes, or river channel networks).

Table 4. F-score metrics for instance segmentation models, considering six percentage thresholds.

Threshold values	**F-score**
70%	0.64
75%	0.65
80%	0.65
85%	0.65
90%	***0.66***
95%	0.58

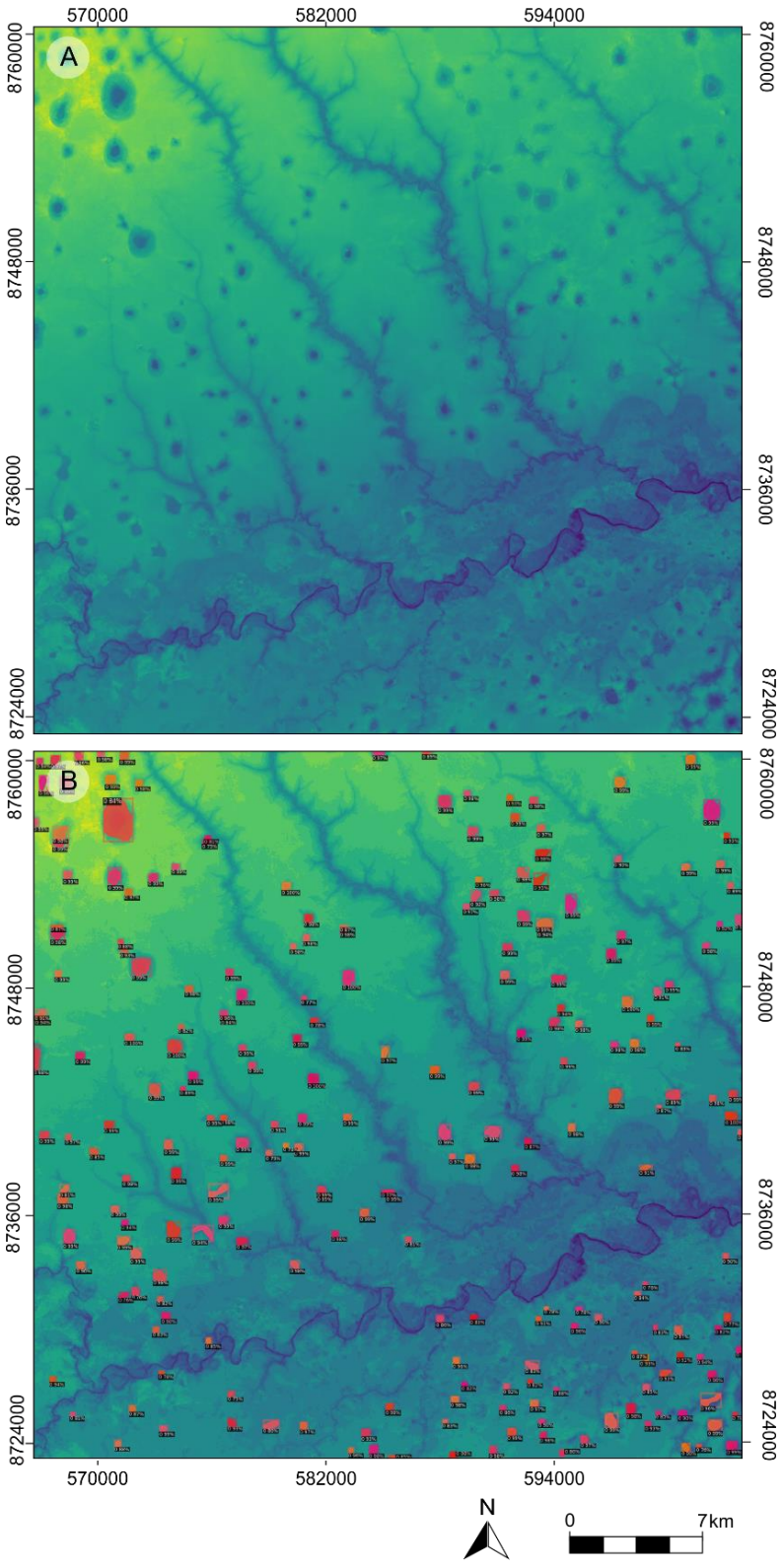

Figure 6. Exemplification of the instance segmentation result after image reconstruction for a subregion of the study area, using Mask-RCNN and a threshold value of 85%.

4.5. Results of Semantic-to-Instance Segmentation Conversion

Semantic-to-instance segmentation conversion allowed generating of instance information from semantic segmentation data. Instance segmentation goes a step further than semantic segmentation, not only assigning class labels but also separating each instance of an object and distinguishing them through unique masks that perform more detailed analysis (object counting, tracking, or instance-specific measurements).

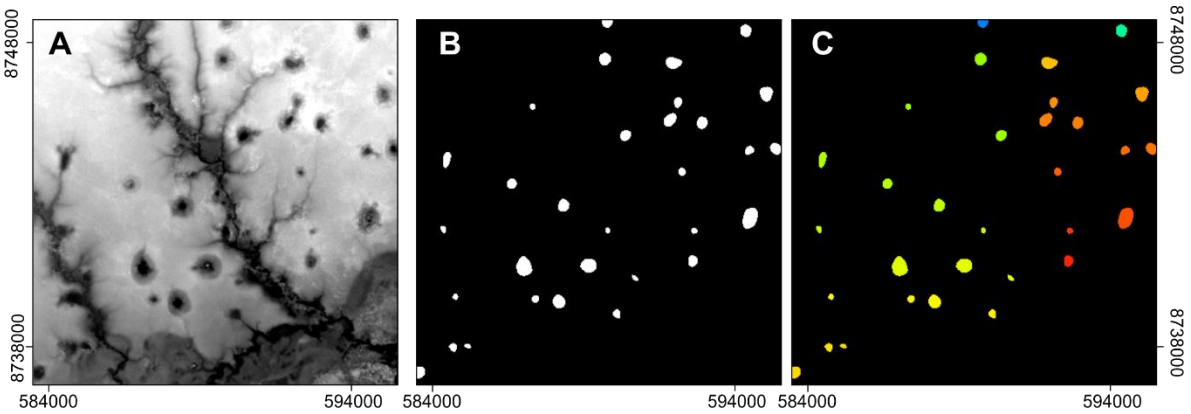

Figure 7. Example of semantic segmentation conversion for instance (A) GLO-30 image; (B) semantic segmentation prediction using FPN of the 11 channels; (C) semantic segmentation-to-instance conversion image, where all karst depressions are given a distinct identifier.

In the conversion process of the studied karst depressions, a crucial factor for the excellent performance was the lack of overlapping or interaction of the instances, representing a more significant challenge in separating individual objects. Therefore, this procedure does not require extensive instance-level annotations and obtains more accurate data using semantic segmentation, especially for small objects, as demonstrated by the accuracy metrics. Semantic segmentation for the best match achieves an F-Score of 0.76 much higher than the 0.66 using Mask-RCNN.

5. DISCUSSION

5.1. Where is the boundary of the karst depressions?

An essential factor to consider in analyzing geomorphological features from deep segmentation methods is the imprecision in defining its limits in labeling reference

samples. The main obstacles in the doline delineation are their irregular geometry, different shapes and sizes, and confusion with other depressions from drainage or river features, forming a more intricate system for detection (Šegina et al., 2018). Therefore, the polygon boundary imprecision that represents the truth is the cause of the accuracy metrics do not reach values as high as other objects with a precise boundary, such as cars and ships. For example, the highest IoU value in semantic segmentation for karst depression was 71.77, considered a low value for other targets with well-defined boundaries. Usually, the delimitation criteria for sinkhole boundaries use algorithms based on a given Digital Terrain Model (Telbisz et al., 2009), such as (a) outermost closed contour line; (b) filling up the doline to the level at which the water overflows from the depression; and (c) polygonal watershed boundaries using interdoline ridges. However, these methods generate different results, where the polygonal watershed boundaries have much larger areas than the other two methods. Besides, these procedures vary significantly with the DEM; using either in deep segmentation would provide a bias in favor of the DEM used. The present study used the visual interpretation of optical images to delimit karst depressions, not favoring any specific DEM.

5.2. Importance of the Karst Depression Dataset

Effective mapping and extracting geomorphological features are fundamental for land management, urban planning, ecological protection, prevention of geological disasters, and management of natural and tourist assets. Feature detection becomes even more relevant in karst areas, where complex topography and ecology make the task more challenging. In this context, DL advances have enabled a new perspective on the ability to classify images, increasing accuracy within a learning process with hierarchical features (from low to high level) directly from the input data. However, mapping geomorphological features from these new techniques needs databases for training, constituting a great challenge and gap in the recognition of complex landscapes. Although there are DL studies in geomorphology, most applications are for mapping landscapes (Buscombe and Goldstein, 2022; Buscombe and Ritchie, 2018; Farmakis-Serebryakova et al., 2022) or geomorphological units (Du et al., 2019; Meij et al., 2022), corresponding to a class of "stuff" formed by amorphous features. Even the study that aims to map the karst cone (Fu et al., 2021) effectively mapped geomorphological units without individualizing the features. Therefore, few studies with DL seek the individualization of

well-defined and countable features, such as karst depressions, which correspond to a class of "things," an instance. Among the mapping of specific features, the detection of yardangs stands out (Farmakis-Serebryakova et al., 2022). Therefore, the growth of DL methods depends on the development of vast training data, and a significant contribution of the present study was the elaboration of a karst depression database.

5.3. Comparison between Digital Elevation Models for detecting karst depressions

The DEM accuracy results in detecting karst depressions are consistent with comparative studies on the vertical accuracy of different global DEMs with a resolution of 30 meters. **Table 5** lists comparative studies that order the DEMs from highest to lowest accuracy value; in red are marked models not used in this research. Although the GLO-30 has few comparative studies because it is newly available, the investigation that uses it proved to be the best result (Guth and Geoffroy, 2021). Like our research, the AW3D30 presents the best result in the absence of GLO-30 and when ignoring other DEMs not used in the present study (LIDAR and TanDEM-X with 12m). In all studies, the worst accuracy results came from the ASTER GDEM.

Table 7. Comparing and ranking the vertical accuracy of Digital Elevation Models based on scientific studies. DEMs not included in the present study are marked in italics: Cartosat-I DEM (CartoDEM), MERIT, LIDAR, Continuous Elevation data for Mexico (CEM), TanDEM-X with 12 m and 90 m resolution.

Ordering DEMs by precision	Study
AW3D30 > SRTM > *CartoDEM* > ASTER GDEM V2	(Jain et al., 2018)
AW3D30 > SRTM > ASTER GDEM	(Florinsky et al., 2018)
TanDEM-X (12 m) > AW3D30 ~ NASADEM ~ *MERIT (90m)* > SRTM > ASTERDEM	(Gesch, 2018)
AW3D30 ~SRTM > ASTER GDEM	(Yap et al., 2019)
AW3D30 > NASADEM ~ SRTM > *TanDEM-X (90 m)* > *MERIT (90 m)* > ASTER GDEM	(Uuemaa et al., 2020).
LIDAR > AW3D30 V3 ~ AW3D30 V2 > NASADEM > SRTM > *CEM* >ASTER GDEM V3 > ASTER GDEM V2	(Carrera-Hernández, 2021)
GLO-30 > AW3D30 ~ SRTM V3~ NASADEM > ASTER GDEM	(Guth and Geoffroy, 2021)
TanDEM-X (12m) > SRTM > ASTER GDEM	(Han et al., 2021)

Few studies performed MED comparisons for the detection of sinkholes. These studies using the DEM-based sink depth technique mostly corroborate the results described: (a) SRTM > AW3D30 > ASTER GDEM (de Carvalho et al., 2013); (b) SRTM ~ASTER GDEM > ALOS PASAR DEM (12.5 meters) (Theilen-Willige, 2018); and (c)

DEM from topographic maps > SRTM > ASTER GDEM > DSM from Sentinel-1 > AW3D30 > DEM from the Greek Cadastral (Kakavas et al., 2018). The last study described presents the most significant difference with the DEM vertical accuracy studies.

5.4. Comparison between semantic segmentation architeture for detecting Karst depressions

The superior performance of the FPN compared to other semantic segmentation models for karst depression detection may be due to its ability to capture multi-scale features by combining high-resolution features from different CNN levels, effectively dealing with variations in size, shape, and context of karst depressions. Therefore, the FPN generates rich intermediary feature maps that capture fine-grained details and high-level semantic information, incorporating feature pyramids and lateral connections, suitable for applications in complex environments such as geomorphic features that often exhibit intricate spatial structures and scale variations. Hence, the effectiveness of CNN architectures in karst relief detection differs from that of clearly defined features like wind farms, which present worse performance for the FPN architecture and better performance for LinkNet, followed by Unet and Unet++. (de Carvalho et al., 2023).

5.5. Analysis of instance segmentation and semantic-to-instance segmentation conversion

Studies show that the use of conversion techniques from semantic segmentation to instance segmentation using GIS software can be a good solution considering the following aspects: (a) creating semantic segmentation datasets is much simpler and less labor intensive than instance segmentation, and (b) instance segmentation methods present inferior performance in the delimitation of small objects as in the present case (de Carvalho et al., 2022c). Besides, the conversion is more straightforward when the objects are naturally isolated, as occurs in karst depressions, not requiring techniques such as inserting edges around the object to ensure isolation between targets that appear in contacts such as vehicles. The simple conversion of raster results into vectors allows discriminating a label for each polygon and its individualization as an instance.

5.6. Implications of mapping karst depressions in environmental analysis

To protect vulnerable karst characteristics, natural resources, and underground water and to prevent economic losses due to subsidence and collapse, the difference in density and distribution of dolines might indicate priority areas for strict conservation and environmental management of human activities developed in areas with karst depressions in Bambuí Group, especially agriculture, urbanization, and infrastructure. Some studies through different techniques using remote sensing data have identified karst depressions in areas of the Bambuí Group (de Carvalho et al., 2013; Ferreira et al., 2022; Guimarães et al., 2005). The spatial distribution of karst depressions indicates that some areas of the Bambuí Group in Bahia are more prone to dissolution due to the high density of dolines, contrasting with other areas with fewer karst depressions. The density of dolines in the Bambuí Group of Bahia was neither indicated in previous studies nor in conservation and environmental management policies. Ferreira et al. (2022) indicate the existence of density patterns of medium to large-size dolines of the Bambuí Group in three identified groups, but the study using semi-automatic GIS-based with DEM data investigated the Corrente river basin, state of Goiás, Brazil. It indicates that the different density distribution of dolines is a primary characteristic of the Bambuí Group karst system.

The study region has a single and reduced protection area in the National Register of Conservation Units, the Environmental Protection Area of Rio Preto (APA Rio Preto), created in 2006 by the state government. This protection category allows private ownership of land and different activities and economic uses, trying to associate production and sustainability, being the most permissive category of protected areas according to Brazilian environmental legislation. The outline of the APA Rio Preto indicates that its creation was to protect the river system and the surrounding native vegetation, disregarding the dolines and the karst system's complexity as fundamental conservation elements. Typically, environmental policies prioritize protecting areas with rich and vulnerable biotic attributes but leave abiotic attributes in the background (Crofts, 2019; Gray, 2019). Therefore, a systemic biotic-abiotic approach is needed (Crofts, 2019) due to the importance of elements and processes of geodiversity for maintaining biodiversity and providing essential services for human societies (Gray, 2019; Tukiainen et al., 2023).

Therefore, the development of environmental policies aimed at the Bambuí Group and the distribution of karst depressions in the landscape is fundamental due to its

environmental fragility. Studies on the Bambuí karst make it possible to identify environmentally sensitive areas and implement different categories of protected areas, from Environmental Protection Areas to more restrictive areas such as National Parks, Biological Reserves, and Ecological Stations.

6. CONCLUSION

In Bambuí Group karst environment, the depressions are the most prevalent topographic features, and the description of their spatial distribution is critical for studying karst aquifers and mitigating damage. The present article innovated in detecting karst depressions using semantic and instance segmentations from deep learning architectures. Instance segmentation has yet to be explored in analyzing the geomorphological targets, allowing the individualization of its occurrences. In addition to using instance segmentation with the Mask-RCNN architecture, we use the sematic-to-instance conversion approach using GIS tools. The research developed a new dataset composed of different DEMs with their geomorphometric attributes and labeled data of karst depressions from the visual interpretation of satellite images (Sentinel-2 and OLI-Landsat 8). Given the resolution of the tested DEMs at 30 meters, the karst depressions are primarily categorized as small objects with dimensions smaller than 32x32 pixels. The results of deep segmentation demonstrated the potential of deep segmentation in the delimitation of karst geomorphological features. FPN architecture using the EfficientNet-B7 backbone obtained the best results among the semantic segmentation architectures, probably due to its capacity to capture multiscale features by integrating high-resolution features from different levels of a CNN. This multiscale feature fusion enables the FPN to effectively handle variations in karst depressions' size, shape, and context. The results with semantic segmentation were superior to the instance segmentation, justifying the implementation of the semantic-to-instance segmentation conversion, which also facilitates the dataset elaboration. Comparing the five DEMs, the GLO-30 obtained the best accuracy results considering the different architectures evaluated, followed by the AW3D30. The worst results were from ASTER GDEM. These results are consistent with other studies on DEMs' vertical accuracy. Therefore, the evaluated methods demonstrate the high detection capacity of these geomorphic features formed in different contexts with high variability of shapes and sizes.

Future studies can apply deep learning segmentation using optical, radar, or LiDAR remote sensing images with high spatial resolution, besides the digital terrain models and their terrain attributes used in this research. Another relevant test would be testing digital terrain models with a higher spatial resolution, like the 12-meter TanDEM-X. The studies can be expanded to other places of karst relief, increasing the available database in Brazil and worldwide.

REFERENCES

Abrams, M., Crippen, R., Fujisada, H., 2020. ASTER Global Digital Elevation Model (GDEM) and ASTER Global Water Body Dataset (ASTWBD). Remote Sens. 12, 1–12. https://doi.org/10.3390/rs12071156

Angel, J.C., Nelson, D.O., Panno, S. V., 2004. Comparison of a new GIS-based technique and a manual method for determining sinkhole density: An example from Illinois' sinkhole plain. J. Cave Karst Stud. 66, 9–17.

Antonić, O., Hatic, D., Pernar, R., 2001. DEM-based depth in sink as an environmental estimator. Ecol. Modell. 138, 247–254. https://doi.org/10.1016/S0304-3800(00)00405-1

Berry, P.A.M., Smith, R.G., Benveniste, J., 2010. ACE2: The New Global Digital Elevation Model, in: Gravity, Geoid and Earth Observation. International Association of Geodesy Symposia. Springer Berlin Heidelberg, pp. 231–237. https://doi.org/10.1007/978-3-642-10634-7_30

Beynen, P.E., 2011. Karst Management. Springer Netherlands, Dordrecht. https://doi.org/10.1007/978-94-007-1207-2

Bhuiyan, M.A.E., Witharana, C., Liljedahl, A.K., Jones, B.M., Daanen, R., Epstein, H.E., Kent, K., Griffin, C.G., Agnew, A., 2020. Understanding the Effects of Optimal Combination of Spectral Bands on Deep Learning Model Predictions: A Case Study Based on Permafrost Tundra Landform Mapping Using High Resolution Multispectral Satellite Imagery. J. Imaging 6, 97. https://doi.org/10.3390/jimaging6090097

Buscombe, D., Goldstein, E.B., 2022. A Reproducible and Reusable Pipeline for Segmentation of Geoscientific Imagery. Earth Sp. Sci. 9, 1–11. https://doi.org/10.1029/2022EA002332

Buscombe, D., Ritchie, A., 2018. Landscape Classification with Deep Neural Networks. Geosciences 8, 244. https://doi.org/10.3390/geosciences8070244

Bystriakova, N., Alves De Melo, P.H., Moat, J., Lughadha, E.N., Monro, A.K., 2019. A Preliminary Evaluation of The Karst Flora of Brazil Using Collections Data. Sci. Rep. 9, 17037. https://doi.org/10.1038/s41598-019-53104-6

Cai, Z., Vasconcelos, N., 2018. Cascade R-CNN: Delving Into High Quality Object Detection, in: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, Salt Lake City, UT, USA, pp. 6154–6162. https://doi.org/10.1109/CVPR.2018.00644

Carrera-Hernández, J.J., 2021. Not all DEMs are equal: An evaluation of six globally available 30 m resolution DEMs with geodetic benchmarks and LiDAR in Mexico. Remote Sens. Environ. 261. https://doi.org/10.1016/j.rse.2021.112474

Carvalho, O.L.F. de, de Carvalho Júnior, O.A., Albuquerque, A.O. de, Bem, P.P. de, Silva, C.R., Ferreira, P.H.G., Moura, R. dos S. de, Gomes, R.A.T., Guimarães, R.F., Borges, D.L., 2021. Instance Segmentation for Large, Multi-Channel Remote Sensing Imagery Using Mask-RCNN and a Mosaicking Approach. Remote Sens. 13, 39. https://doi.org/10.3390/rs13010039

Carvalho Júnior, O.A., Berbet-Born, M., Martins, E.S., Guimarães, R.F., Gomes, R.A.T., 2008. Ambientes Cársticos, in: Florenzano, T.G. (Ed.), Geomorfologia: Conceitos e Tecnologias Atuais. Oficina de Textos, São Paulo, pp. 183–218.

Catani, F., 2021. Landslide detection by deep learning of non-nadiral and crowdsourced optical images. Landslides 18, 1025–1044. https://doi.org/10.1007/s10346-020-01513-4

Chaurasia, A., Culurciello, E., 2017. LinkNet: Exploiting encoder representations for efficient semantic segmentation, in: 2017 IEEE Visual Communications and Image Processing (VCIP). IEEE, St. Petersburg, FL, USA, pp. 1–4. https://doi.org/10.1109/VCIP.2017.8305148

Chen, K., Ouyang, W., Loy, C.C., Lin, D., Pang, J., Wang, J., Xiong, Y., Li, X., Sun, S., Feng, W., Liu, Z., Shi, J., 2019. Hybrid Task Cascade for Instance Segmentation, in: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Long Beach, CA, USA, USA, pp. 4969–4978. https://doi.org/10.1109/CVPR.2019.00511

Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L., 2018. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. IEEE Trans. Pattern Anal. Mach. Intell. 40, 834–848. https://doi.org/10.1109/TPAMI.2017.2699184

Ciglič, R., Čonč, Š., Breg Valjavec, M., 2022. The Impact of Digital Elevation Model

Preprocessing and Detection Methods on Karst Depression Mapping in Densely Forested Dinaric Mountains. Remote Sens. 14, 2416. https://doi.org/10.3390/rs14102416

Closson, D., Patil, A., Musthafa, M., Gallagher, M., Das, N., 2023. Impact of Dead Sea Halo-Karst Development on an Earthen Dike Rehabilitation Project. Geosci. 13. https://doi.org/10.3390/geosciences13020042

Coelho, J.C.C., Martins-Neto, M.A., Marinho, M.S., 2008. Estilos estruturais e evolução tectônica da porção mineira da bacia proterozóica do São Francisco. Rev. Bras. Geociências 38, 149–165. https://doi.org/10.25249/0375-7536.2008382S149165

Crippen, R., Buckley, S., Agram, P., Belz, E., Gurrola, E., Hensley, S., Kobrick, M., Lavalle, M., Martin, J., Neumann, M., Nguyen, Q., Rosen, P., Shimada, J., Simard, M., Tung, W., 2016. NASADEM GLOBAL ELEVATION MODEL: METHODS AND PROGRESS. ISPRS - Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XLI-B4, 125–128. https://doi.org/10.5194/isprsarchives-XLI-B4-125-2016

Crofts, R., 2019. Linking geoconservation with biodiversity conservation in protected areas. Int. J. Geoheritage Park. 7, 211–217. https://doi.org/10.1016/j.ijgeop.2019.12.002

Cuellar, A.C., Cenci, L., Santella, C., Albinet, C., 2022. Evaluating the Copernicus Dem Dataset Potential for the Identification of (Flash) Flood-Prone Areas by Using a Geomorphological Approach. Int. Geosci. Remote Sens. Symp. 2022-July, 5997–6000. https://doi.org/10.1109/IGARSS46834.2022.9884948

de Albuquerque, A.O., de Carvalho Júnior, O.A., Carvalho, O.L.F. de, de Bem, P.P., Ferreira, P.H.G., de Moura, R. dos S., Silva, C.R., Trancoso Gomes, R.A., Fontes Guimarães, R., 2020. Deep Semantic Segmentation of Center Pivot Irrigation Systems from Remotely Sensed Data. Remote Sens. 12, 2159. https://doi.org/10.3390/rs12132159

de Carvalho, O., Guimarães, R., Montgomery, D., Gillespie, A., Trancoso Gomes, R., de Souza Martins, É., Silva, N., 2013. Karst Depression Detection Using ASTER, ALOS/PRISM and SRTM-Derived Digital Elevation Models in the Bambuí Group, Brazil. Remote Sens. 6, 330–351. https://doi.org/10.3390/rs6010330

de Carvalho, O.L.F., de Carvalho Junior, O.A., de Albuquerque, A.O., Orlandi, A.G., Hirata, I., Borges, D.L., Gomes, R.A.T., Guimarães, R.F., 2023. A Data-Centric Approach for Wind Plant Instance-Level Segmentation Using Semantic Segmentation and GIS. Remote Sens. 15, 1240. https://doi.org/10.3390/rs15051240

de Carvalho, O.L.F., de Carvalho Júnior, O.A., de Albuquerque, A.O., Santana, N.C., Borges, D.L., Luiz, A.S., Gomes, R.A.T., Guimarães, R.F., 2022a. Multispectral panoptic segmentation: Exploring the beach setting with worldview-3 imagery. Int. J. Appl. Earth Obs. Geoinf. 112, 102910. https://doi.org/10.1016/j.jag.2022.102910

de Carvalho, O.L.F., de Carvalho Júnior, O.A., Silva, C.R. e, de Albuquerque, A.O., Santana, N.C., Borges, D.L., Gomes, R.A.T., Guimarães, R.F., 2022b. Panoptic Segmentation Meets Remote Sensing. Remote Sens. 14, 965. https://doi.org/10.3390/rs14040965

de Carvalho, O.L.F., Junior, O.A. de C., de Albuquerque, A.O., Santana, N.C., Guimaraes, R.F., Gomes, R.A.T., Borges, D.L., 2022c. Bounding Box-Free Instance Segmentation Using Semi-Supervised Iterative Learning for Vehicle Detection. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 15, 3403–3420. https://doi.org/10.1109/JSTARS.2022.3169128

de Castro Tayer, T., Rodrigues, P.C.H., 2021. Assessment of a semi-automatic spatial analysis method to identify and map sinkholes in the Carste Lagoa Santa environmental protection unit, Brazil. Environ. Earth Sci. 80, 83. https://doi.org/10.1007/s12665-020-09354-z

de Oliveira, S.N., de Carvalho Júnior, O.A., Gomes, R.A.T., Guimarães, R.F., McManus, C.M., 2017a. Landscape-fragmentation change due to recent agricultural expansion in the Brazilian Savanna, Western Bahia, Brazil. Reg. Environ. Chang. 17, 411–423. https://doi.org/10.1007/s10113-016-0960-0

de Oliveira, S.N., de Carvalho Júnior, O.A., Trancoso Gomes, R.A., Fontes Guimarães, R., McManus, C.M., 2017b. Deforestation analysis in protected areas and scenario simulation for structural corridors in the agricultural frontier of Western Bahia, Brazil. Land use policy 61, 40–52. https://doi.org/10.1016/j.landusepol.2016.10.046

De Waele, J., Gutiérrez, F., Parise, M., Plan, L., 2011. Geomorphology and natural hazards in karst areas: A review. Geomorphology 134, 1–8. https://doi.org/10.1016/j.geomorph.2011.08.001

Doctor, D., Young, J., 2013. An Evaluation of Automated GIS Tools for Delineating Karst Sinkholes and Closed Depressions from 1-Meter LiDAR-Derived Digital Elevation Data, in: Full Proceedings of the Thirteenth Multidisciplinary Conference on Sinkholes and the Engineering and Environmental Impacts of Karst. National Cave and Karst Research Institute, Carlsbad, New Mexico, USA, pp. 449–458. https://doi.org/10.5038/9780979542275.1156

Dou, J., Li, X., Yunus, A.P., Paudel, U., Chang, K.-T., Zhu, Z., Pourghasemi, H.R., 2015. Automatic detection of sinkhole collapses at finer resolutions using a multi-component remote

sensing approach. Nat. Hazards 78, 1021–1044. https://doi.org/10.1007/s11069-015-1756-0

Du, H., Wang, J., Han, C., 2022. High-precision remote sensing mapping of aeolian sand landforms based on deep learning algorithms. Open Geosci. 14, 224–233. https://doi.org/10.1515/geo-2022-0351

Du, L., You, X., Li, K., Meng, L., Cheng, G., Xiong, L., Wang, G., 2019. Multi-modal deep learning for landform recognition. ISPRS J. Photogramm. Remote Sens. 158, 63–75. https://doi.org/10.1016/j.isprsjprs.2019.09.018

Farmakis-Serebryakova, M., Heitzler, M., Hurni, L., 2022. Terrain Segmentation Using a U-Net for Improved Relief Shading. ISPRS Int. J. Geo-Information 11, 395. https://doi.org/10.3390/ijgi11070395

Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., Alsdorf, D., 2007. The Shuttle Radar Topography Mission. Rev. Geophys. 45, RG2004. https://doi.org/10.1029/2005RG000183

Ferranti, J., 2021. Digital elevation data [WWW Document]. URL http://viewfinderpanoramas.org/dem3.html (accessed 1.20.23).

Ferreira, C.F., Hussain, Y., Uagoda, R., 2022. a Semi-Automatic Approach for Doline Mapping in Brazilian Covered Karst: the Way Forward To Vulnerability Assessment. Acta Carsologica 51, 19–31. https://doi.org/10.3986/ac.v51i1.10011

Fleury, S., 2009. Land Use Policy and Practice on Karst Terrains. Springer Netherlands, Dordrecht. https://doi.org/10.1007/978-1-4020-9670-9

Florinsky, I. V., Skrypitsyna, T.N., Luschikova, O.S., 2018. Comparative accuracy of the AW3D30 DSM, ASTER GDEM, AND SRTM1 DEM: A case study on the Zaoksky testing ground, central European Russia. Remote Sens. Lett. 9, 706–714. https://doi.org/10.1080/2150704X.2018.1468098

Ford, D., Williams, P., 2007. Karst Hydrogeology and Geomorphology. John Wiley & Sons Ltd,., West Sussex, England. https://doi.org/10.1002/9781118684986

Fu, H., Fu, B., Shi, P., 2021. An Improved Segmentation Method for Automatic Mapping of Cone Karst from Remote Sensing Data Based on DeepLab V3+ Model. Remote Sens. 13, 441.

https://doi.org/10.3390/rs13030441

Gao, B., Chen, N., Blaschke, T., Wu, C.Q., Chen, J., Xu, Y., Yang, X., Du, Z., 2021. Automated Characterization of Yardangs Using Deep Convolutional Neural Networks. Remote Sens. 13, 733. https://doi.org/10.3390/rs13040733

Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V., Martinez-Gonzalez, P., Garcia-Rodriguez, J., 2018. A survey on deep learning techniques for image and video semantic segmentation. Appl. Soft Comput. 70, 41–65. https://doi.org/10.1016/j.asoc.2018.05.018

Geng, Q., Zhou, Z., Cao, X., 2018. Survey of recent progress in semantic image segmentation with CNNs. Sci. China Inf. Sci. 61, 1–18. https://doi.org/10.1007/s11432-017-9189-6

Gesch, D.B., 2018. Best Practices for Elevation-Based Assessments of Sea-Level Rise and Coastal Flooding Exposure. Front. Earth Sci. 6. https://doi.org/10.3389/feart.2018.00230

Goldscheider, N., Chen, Z., Auler, A.S., Bakalowicz, M., Broda, S., Drew, D., Hartmann, J., Jiang, G., Moosdorf, N., Stevanovic, Z., Veni, G., 2020. Global distribution of carbonate rocks and karst water resources. Hydrogeol. J. 28, 1661–1677. https://doi.org/10.1007/s10040-020-02139-5

Gray, M., 2019. Geodiversity, geoheritage and geoconservation for society. Int. J. Geoheritage Park. 7, 226–236. https://doi.org/10.1016/j.ijgeop.2019.11.001

Grohmann, C.H., 2018. Evaluation of TanDEM-X DEMs on selected Brazilian sites: Comparison with SRTM, ASTER GDEM and ALOS AW3D30. Remote Sens. Environ. 212, 121–133. https://doi.org/10.1016/j.rse.2018.04.043

Gu, W., Bai, S., Kong, L., 2022. A review on 2D instance segmentation based on deep neural networks. Image Vis. Comput. 120, 104401. https://doi.org/10.1016/j.imavis.2022.104401

Guimarães, R.F., de Carvalho Júnior, O.A., de Souza Martins, E., Ferreira de Carvalho, A.P., Trancoso Gomes, R.A., 2005. Detection of karst depression by aster image in the Bambui Group, Brazil, in: Ehlers, M., Michel, U. (Eds.), Proc. SPIE 5983, Remote Sensing for Environmental Monitoring, GIS Applications, and Geology V. Bruges, Belgium, p. 59831H. https://doi.org/10.1117/12.627741

Guo, Y., Liu, Y., Georgiou, T., Lew, M.S., 2018. A review of semantic segmentation using deep neural networks. Int. J. Multimed. Inf. Retr. 7, 87–93. https://doi.org/10.1007/s13735-017-0141-

z

Gurgel, R.S., Carvalho Júnior, O.A. de, Gomes, R.A.T., Guimarães, R.F., Martins, É.D.S., 2013. Relação entre a evolução do uso da terra com as unidades geomorfológicas no município de Riachão das Neves (BA). GeoTextos 9, 177–201. https://doi.org/10.9771/1984-5537geo.v9i1.6477

Guth, P.L., Geoffroy, T.M., 2021. LiDAR point cloud and ICESat-2 evaluation of 1 second global digital elevation models: Copernicus wins. Trans. GIS 25, 2245–2261. https://doi.org/10.1111/tgis.12825

Gutiérrez, F., Parise, M., De Waele, J., Jourde, H., 2014. A review on natural and human-induced geohazards and impacts in karst. Earth-Science Rev. 138, 61–88. https://doi.org/10.1016/j.earscirev.2014.08.002

Hafiz, A.M., Bhat, G.M., 2020. A survey on instance segmentation: state of the art. Int. J. Multimed. Inf. Retr. 9, 171–189. https://doi.org/10.1007/s13735-020-00195-x

Han, H., Zeng, Q., Jiao, J., 2021. Quality Assessment of TanDEM-X DEMs, SRTM and ASTER GDEM on Selected Chinese Sites. Remote Sens. 13, 1304. https://doi.org/10.3390/rs13071304

Hawker, L., Uhe, P., Paulo, L., Sosa, J., Savage, J., Sampson, C., Neal, J., 2022. A 30 m global map of elevation with forests and buildings removed. Environ. Res. Lett. 17, 024016. https://doi.org/10.1088/1748-9326/ac4d4f

He, K., Gkioxari, G., Dollar, P., Girshick, R., 2020. Mask R-CNN. IEEE Trans. Pattern Anal. Mach. Intell. 42, 386–397. https://doi.org/10.1109/TPAMI.2018.2844175

Hofierka, J., Gallay, M., Bandura, P., Šašak, J., 2018. Identification of karst sinkholes in a forested karst landscape using airborne laser scanning data and water flow analysis. Geomorphology 308, 265–277. https://doi.org/10.1016/j.geomorph.2018.02.004

Huang, L., Lantz, T.C., Fraser, R.H., Tiampo, K.F., Willis, M.J., Schaefer, K., 2022. Accuracy, Efficiency, and Transferability of a Deep Learning Model for Mapping Retrogressive Thaw Slumps across the Canadian Arctic. Remote Sens. 14, 2747. https://doi.org/10.3390/rs14122747

Huang, L., Liu, L., Jiang, L., Zhang, T., 2018. Automatic Mapping of Thermokarst Landforms from Remote Sensing Images Using Deep Learning: A Case Study in the Northeastern Tibetan Plateau. Remote Sens. 10, 2067. https://doi.org/10.3390/rs10122067

Huang, Z., Huang, L., Gong, Y., Huang, C., Wang, X., 2019. Mask Scoring R-CNN, in: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Long Beach, CA, USA, USA, pp. 6402–6411. https://doi.org/10.1109/CVPR.2019.00657

Jain, A.O., Thaker, T., Chaurasia, A., Patel, P., Singh, A.K., 2018. Vertical accuracy evaluation of SRTM-GL1, GDEM-V2, AW3D30 and CartoDEM-V3.1 of 30-m resolution with dual frequency GNSS for lower Tapi Basin India, Geocarto International. Taylor & Francis. https://doi.org/10.1080/10106049.2017.1343392

Jiang, S., Wu, F., Yung, K.L., Yang, Y., Ip, W.H., Gao, M., Foster, J.A., 2021. A robust end-to-end deep learning framework for detecting Martian landforms with arbitrary orientations. Knowledge-Based Syst. 234, 107562. https://doi.org/10.1016/j.knosys.2021.107562

Kakavas, M., Nikolakopoulos, K.G., Kyriou, A., Zagana, H., 2018. Assessment of freely available DSMs for automatic karst feature detection. Arab. J. Geosci. 11, 388. https://doi.org/10.1007/s12517-018-3654-8

Karmann, I., Sánchez, L.E., 1979. Distribuição das rochas carbonáticas e províncias espeleológicas do Brasil. Espeleo-Tema 13, 105–167.

Kim, Y.J., Nam, B.H., Youn, H., 2019. Sinkhole Detection and Characterization Using LiDAR-Derived DEM with Logistic Regression. Remote Sens. 11, 1592. https://doi.org/10.3390/rs11131592

Kobal, M., Bertoncelj, I., Pirotti, F., Dakskobler, I., Kutnar, L., 2015. Using lidar data to analyse sinkhole characteristics relevant for understory vegetation under forest cover - Case study of a high karst area in the Dinaric mountains. PLoS One 10, 1–19. https://doi.org/10.1371/journal.pone.0122070

Kuchenbecker, M., Reis, H.L.S., Silva, L.C. da, Costa, R.D. da, Fragoso, D.G.C., Knauer, L.G., Dussin, I.A., Soares, A.C.P., 2015. Age constraints for deposition and sedimentary provenance of Espinhaço Supergroup and Bambuí Group in eastern São Francisco Craton. Geonomos. https://doi.org/10.18285/geonomos.v23i2.708

Lateef, F., Ruichek, Y., 2019. Survey on semantic segmentation using deep learning techniques. Neurocomputing 338, 321–348. https://doi.org/10.1016/j.neucom.2019.02.003

Lee, E.J., Shin, S.Y., Ko, B.C., Chang, C., 2016. Early sinkhole detection using a drone-based thermal camera and image processing. Infrared Phys. Technol. 78, 223–232.

https://doi.org/10.1016/j.infrared.2016.08.009

Li, S., Xiong, L., Tang, G., Strobl, J., 2020. Deep learning-based approach for landform classification from integrated data sources of digital elevation model and imagery. Geomorphology 354, 107045. https://doi.org/10.1016/j.geomorph.2020.107045

Li, W., Hsu, C.-Y., 2020. Automated terrain feature identification from remote sensing imagery: a deep learning approach. Int. J. Geogr. Inf. Sci. 34, 637–660. https://doi.org/10.1080/13658816.2018.1542697

Li, X., Chen, D., 2022. A survey on deep learning-based panoptic segmentation. Digit. Signal Process. 120, 103283. https://doi.org/10.1016/j.dsp.2021.103283

Li, Y., Qi, H., Dai, J., Ji, X., Wei, Y., 2017. Fully Convolutional Instance-Aware Semantic Segmentation, in: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Honolulu, HI, USA, pp. 4438–4446. https://doi.org/10.1109/CVPR.2017.472

Lin, T.-Y., Dollar, P., Girshick, R., He, K., Hariharan, B., Belongie, S., 2017. Feature Pyramid Networks for Object Detection, in: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Honolulu, HI, USA, pp. 936–944. https://doi.org/10.1109/CVPR.2017.106

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., Zitnick, C.L., 2014. Microsoft COCO: Common Objects in Context, in: Fleet, D., Tomas, P., Schiele, B., Tuytelaars, T. (Eds.), Computer Vision – ECCV 2014. Lecture Notes in Computer Science, Vol 8693. Springer, Chan, pp. 740–755. https://doi.org/10.1007/978-3-319-10602-1_48

Martins-Neto, M.., Pedrosa-Soares, A.., Lima, S.A.., 2001. Tectono-sedimentary evolution of sedimentary basins from Late Paleoproterozoic to Late Neoproterozoic in the São Francisco craton and Araçuaí fold belt, eastern Brazil. Sediment. Geol. 141–142, 343–370. https://doi.org/10.1016/S0037-0738(01)00082-3

Martins-Neto, M.A., 2009. Sequence stratigraphic framework of Proterozoic successions in eastern Brazil. Mar. Pet. Geol. 26, 163–176. https://doi.org/10.1016/j.marpetgeo.2007.10.001

Masilela, M., Beckedahl, H., 2022. Karst geomorphology and related environmental problems in Southern Africa – A review. J. African Earth Sci. 196, 104686. https://doi.org/10.1016/j.jafrearsci.2022.104686

Meij, W.M., Meijles, E.W., Marcos, D., Harkema, T.T.L., Candel, J.H.J., Maas, G.J., 2022. Comparing geomorphological maps made manually and by deep learning. Earth Surf. Process. Landforms 47, 1089–1107. https://doi.org/10.1002/esp.5305

Miao, X., Qiu, X., Wu, S.-S., Luo, J., Gouzie, D.R., Xie, H., 2013. Developing Efficient Procedures for Automated Sinkhole Extraction from Lidar DEMs. Photogramm. Eng. Remote Sens. 79, 545–554. https://doi.org/10.14358/PERS.79.6.545

Moreno-Gómez, M., Liedl, R., Stefan, C., 2019. A New GIS-Based Model for Karst Dolines Mapping Using LiDAR; Application of a Multidepth Threshold Approach in the Yucatan Karst, Mexico. Remote Sens. 11, 1147. https://doi.org/10.3390/rs11101147

Mou, L., Zhu, X.X., 2018. Vehicle Instance Segmentation From Aerial Image and Video Using a Multitask Learning Residual Fully Convolutional Network. IEEE Trans. Geosci. Remote Sens. 56, 6699–6711. https://doi.org/10.1109/TGRS.2018.2841808

Na, J., Xue, K., Xiong, L., Tang, G., Ding, H., Strobl, J., Pfeifer, N., 2020. UAV-Based Terrain Modeling under Vegetation in the Chinese Loess Plateau: A Deep Learning and Terrain Correction Ensemble Framework. Remote Sens. 12, 3318. https://doi.org/10.3390/rs12203318

Obu, J., Podobnikar, T., 2013. Algorithm for karst depression recognition using digital terrain models. Geod. Vestn. 57, 260–270. https://doi.org/10.15292/geodetski-vestnik.2013.02.260-270

Palafox, L.F., Hamilton, C.W., Scheidt, S.P., Alvarez, A.M., 2017. Automated detection of geological landforms on Mars using Convolutional Neural Networks. Comput. Geosci. 101, 48–56. https://doi.org/10.1016/j.cageo.2016.12.015

Parise, M., Closson, D., Gutiérrez, F., Stevanović, Z., 2015. Anticipating and managing engineering problems in the complex karst environment. Environ. Earth Sci. 74, 7823–7835. https://doi.org/10.1007/s12665-015-4647-5

Parise, M., Pisano, L., Zumpano, V., 2020. Detection and Characterization of Sinkholes Through Integration of Field Surveys and Semi-automated Techniques, in: Bertrand, C., Denimal, S., Steinmann, M., Renard, P. (Eds.), Eurokarst 2018. Springer, Cham, Besançon, pp. 3–9. https://doi.org/10.1007/978-3-030-14015-1_1

Pasquetti, F., Bini, M., Ciampalini, A., 2019. Accuracy of the TanDEM-X Digital Elevation Model for Coastal Geomorphological Studies in Patagonia (South Argentina). Remote Sens. 11, 1767. https://doi.org/10.3390/rs11151767

Paula-Santos, G.M., Babinski, M., Kuchenbecker, M., Caetano-Filho, S., Trindade, R.I., Pedrosa-Soares, A.C., 2015. New evidence of an Ediacaran age for the Bambuí Group in southern São Francisco craton (eastern Brazil) from zircon U–Pb data and isotope chemostratigraphy. Gondwana Res. 28, 702–720. https://doi.org/10.1016/j.gr.2014.07.012

Pimentel, M.M., Rodrigues, J.B., DellaGiustina, M.E.S., Junges, S., Matteini, M., Armstrong, R., 2011. The tectonic evolution of the Neoproterozoic Brasília Belt, central Brazil, based on SHRIMP and LA-ICPMS U–Pb sedimentary provenance data: A review. J. South Am. Earth Sci. 31, 345–357. https://doi.org/10.1016/j.jsames.2011.02.011

Rabus, B., Eineder, M., Roth, A., Bamler, R., 2003. The shuttle radar topography mission—a new class of digital elevation models acquired by spaceborne radar. ISPRS J. Photogramm. Remote Sens. 57, 241–262. https://doi.org/10.1016/S0924-2716(02)00124-7

Rafique, M.U., Zhu, J., Jacobs, N., 2022. Automatic Segmentation of Sinkholes Using a Convolutional Neural Network. Earth Sp. Sci. 9, 1–15. https://doi.org/10.1029/2021EA002195

Reuter, H.I., Nelson, A., Jarvis, A., 2007. An evaluation of void-filling interpolation methods for SRTM data. Int. J. Geogr. Inf. Sci. 21, 983–1008. https://doi.org/10.1080/13658810601169899

Ribeiro, J.F., Walter, B.M.T., 2002. As Principais Fitofisionomias do Bioma Cerrado, in: Sano, S.M., Almeida, S.P. de, Ribeiro, J.F. (Eds.), Cerrado Ecologia e Flora. Embrapa Cerrados, Brasília, DF, pp. 151–199.

Robinson, N., Regetz, J., Guralnick, R.P., 2014. EarthEnv-DEM90: A nearly-global, void-free, multi-scale smoothed, 90m digital elevation model from fused ASTER and SRTM data. ISPRS J. Photogramm. Remote Sens. 87, 57–67. https://doi.org/10.1016/j.isprsjprs.2013.11.002

Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation, in: Navab, N., Hornegger, J., Wells, W., Frangi, A. (Eds.), Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer, Cham, pp. 234–241. https://doi.org/10.1007/978-3-319-24574-4_28

Santos, R.V., de Alvarenga, C.J.S., Dardenne, M.A., Sial, A.N., Ferreira, V.P., 2000. Carbon and oxygen isotope profiles across Meso-Neoproterozoic limestones from central Brazil: Bambuí and Paranoá groups. Precambrian Res. 104, 107–122. https://doi.org/10.1016/S0301-9268(00)00082-6

Seale, L.D., Florea, L.J., Vacher, H.L., Brinkmann, R., 2008. Using ALSM to map sinkholes in

the urbanized covered karst of Pinellas County, Florida - 1, Methodological considerations. Environ. Geol. 54, 995–1005. https://doi.org/10.1007/s00254-007-0890-8

Šegina, E., Benac, Č., Rubinić, J., Knez, M., 2018. Morphometric analyses of dolines — the problem of delineation and calculation of basic parameters. Acta Carsologica 47. https://doi.org/10.3986/ac.v47i1.4941

Shumack, S., Hesse, P., Farebrother, W., 2020. Deep learning for dune pattern mapping with the AW3D30 global surface model. Earth Surf. Process. Landforms 45, 2417–2431. https://doi.org/10.1002/esp.4888

Sial, A.N., Dardenne, M.A., Misi, A., Pedreira, A.J., Gaucher, C., Ferreira, V.P., Silva Filho, M.A., Uhlein, A., Pedrosa-Soares, A.C., Santos, R.V., Egydio-Silva, M., Babinski, M., Alvarenga, C.J.S., Fairchild, T.R., Pimentel, M.M., 2009. The São Francisco Palaeocontinent, in: Gaucher, Claudio, Sial, Alcides N., Frimmel, H.E., Halverson, G.P. (Eds.), Neoproterozoic-Cambrian Tectonics, Global Change And Evolution: A Focus On South Western Gondwana. pp. 31–69. https://doi.org/10.1016/S0166-2635(09)01603-X

Siart, C., Bubenzer, O., Eitel, B., 2009. Combining digital elevation data (SRTM/ASTER), high resolution satellite imagery (Quickbird) and GIS for geomorphological mapping: A multi-component case study on Mediterranean karst in Central Crete. Geomorphology 112, 106–121. https://doi.org/10.1016/j.geomorph.2009.05.010

Slater, J.A., Garvey, G., Johnston, C., Haase, J., Heady, B., Kroenung, G., Little, J., 2006. The SRTM data “finishing” process and products. Photogramm. Eng. Remote Sensing 72, 237–247. https://doi.org/10.14358/PERS.72.3.237

Su, H., Wei, S., Liu, S., Liang, J., Wang, C., Shi, J., Zhang, X., 2020. HQ-ISNet: High-Quality Instance Segmentation for Remote Sensing Imagery. Remote Sens. 12, 989. https://doi.org/10.3390/rs12060989

Tachikawa, T., Hato, M., Kaku, M., Iwasaki, A., 2011. Characteristics of ASTER GDEM version 2, in: 2011 IEEE International Geoscience and Remote Sensing Symposium. IEEE, pp. 3657–3660. https://doi.org/10.1109/IGARSS.2011.6050017

Tadono, T., Takaku, J., Tsutsui, K., Oda, F., Nagai, H., 2015. Status of “ALOS World 3D (AW3D)” global DSM generation, in: 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). IEEE, Milan, Italy, pp. 3822–3825. https://doi.org/10.1109/IGARSS.2015.7326657

Takaku, J., Tadono, T., Tsutsui, K., Ichikawa, M., 2018. Quality Improvements of 'AW3D' Global Dsm Derived from Alos Prism, in: IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium. IEEE, Valencia, Spain, pp. 1612–1615. https://doi.org/10.1109/IGARSS.2018.8518360

Tan, M., Le, Q. V., 2019. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks, in: Proceedings of the 36th International Conference on Machine Learning. Long Beach, California, USA, pp. 6105–6114.

Telbisz, T., Dragušica, H., Nagy, B., 2009. Doline Morphometric Analysis and Karst Morphology of Biokovo Mt (Croatia) Based on Field Observations and Digital Terrain Analysis. Hrvat. Geogr. Glas. Geogr. Bull. 71, 2–22. https://doi.org/10.21861/HGG.2009.71.02.01

Telbisz, T., Látos, T., Deák, M., Székely, B., Koma, Z., Standovár, T., 2016. The advantage of lidar digital terrain models in doline morphometry compared to topographic map based datasets – Aggtelek karst (Hungary) as an example. Acta Carsologica 45, 5–18. https://doi.org/10.3986/ac.v45i1.4138

Theilen-Willige, B., 2018. Detection of Karst Features in the Black Hills Area in South Dakota/Wyoming, USA, Based on Evaluations of Remote Sensing Data. Geosciences 8, 192. https://doi.org/10.3390/geosciences8060192

Tian, D., Han, Y., Wang, B., Guan, T., Gu, H., Wei, W., 2021. Review of object instance segmentation based on deep learning. J. Electron. Imaging 31, 041205. https://doi.org/10.1117/1.JEI.31.4.041205

Tong, K., Wu, Y., Zhou, F., 2020. Recent advances in small object detection based on deep learning: A review. Image Vis. Comput. 97, 103910. https://doi.org/10.1016/j.imavis.2020.103910

Tukiainen, H., Toivanen, M., Maliniemi, T., 2023. Geodiversity and Biodiversity. Geol. Soc. London, Spec. Publ. 530. https://doi.org/10.1144/SP530-2022-107

Uuemaa, E., Ahi, S., Montibeller, B., Muru, M., Kmoch, A., 2020. Vertical accuracy of freely available global digital elevation models (Aster, aw3d30, merit, tandem-x, srtm, and nasadem). Remote Sens. 12, 1–23. https://doi.org/10.3390/rs12213482

Vagen, T.-G., 2010. Africa Soil Information Service: Hydrologically Corrected/Adjusted SRTM DEM (AfrHySRTM) [WWW Document]. URL

https://cmr.earthdata.nasa.gov/search/concepts/C1214155420-SCIOPS (accessed 1.25.23).

Vu, H.N., Pham, C., Dung, N.M., Ro, S., 2020. Detecting and Tracking Sinkholes Using Multi-Level Convolutional Neural Networks and Data Association. IEEE Access 8, 132625–132641. https://doi.org/10.1109/ACCESS.2020.3010885

Wall, J., Bohnenstiehl, D.R., Wegmann, K.W., Levine, N.S., 2017. Morphometric comparisons between automated and manual karst depression inventories in Apalachicola National Forest, Florida, and Mammoth Cave National Park, Kentucky, USA. Nat. Hazards 85, 729–749. https://doi.org/10.1007/s11069-016-2600-x

Warren, L.V., Quaglio, F., Riccomini, C., Simões, M.G., Poiré, D.G., Strikis, N.M., Anelli, L.E., Strikis, P.C., 2014. The puzzle assembled: Ediacaran guide fossil Cloudina reveals an old proto-Gondwana seaway. Geology 42, 391–394. https://doi.org/10.1130/G35304.1

Weishampel, J., Hightower, J., Chase, A., Chase, D., Patrick, R., 2011. Detection and Morphologic Analysis of Potential Below-Canopy Cave Openings in the Karst Landscape around the Maya Polity of Caracol using Airborne Lidar. J. Cave Karst Stud. 73, 187–196. https://doi.org/10.4311/2010EX0179R1

Wood, J., 1996. The geomorphological characterisation of digital elevation models. University of Leicester.

Wright, J., Barrett, A.M., Fawdon, P., Favaro, E.A., Balme, M.R., Woods, M.J., Karachalios, S., 2022. Jezero crater, Mars: application of the deep learning NOAH-H terrain classification system. J. Maps 18, 484–496. https://doi.org/10.1080/17445647.2022.2095935

Wu, Q., Deng, C., Chen, Z., 2016. Automated delineation of karst sinkholes from LiDAR-derived digital elevation models. Geomorphology 266, 1–10. https://doi.org/10.1016/j.geomorph.2016.05.006

Wu, Q., Liu, H., Wang, S., Yu, B., Beck, R., Hinkel, K., 2015. A localized contour tree method for deriving geometric and topological properties of complex surface depressions based on high-resolution topographical data. Int. J. Geogr. Inf. Sci. 29, 2041–2060. https://doi.org/10.1080/13658816.2015.1038719

Wu, X., Yan, S., Qi, J., Zeng, H., 2020. Deep Learning for Characterizing Paleokarst Collapse Features in 3-D Seismic Images. J. Geophys. Res. Solid Earth 125, 1–23. https://doi.org/10.1029/2020JB019685

Xiong, L., Tang, G., Yang, X., Li, F., 2021. Geomorphology-oriented digital terrain analysis: Progress and perspectives. J. Geogr. Sci. 31, 456–476. https://doi.org/10.1007/s11442-021-1853-9

Xu, Y., Zhu, H., Hu, C., Liu, H., Cheng, Y., 2022. Deep learning of DEM image texture for landform classification in the Shandong area, China. Front. Earth Sci. 16, 352–367. https://doi.org/10.1007/s11707-021-0884-y

Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., O'Loughlin, F., Neal, J.C., Sampson, C.C., Kanae, S., Bates, P.D., 2017. A high-accuracy map of global terrain elevations. Geophys. Res. Lett. 44, 5844–5853. https://doi.org/10.1002/2017GL072874

Yap, L., Kandé, L.H., Nouayou, R., Kamguia, J., Ngouh, N.A., Makuate, M.B., 2019. Vertical accuracy evaluation of freely available latest high-resolution (30 m) global digital elevation models over Cameroon (Central Africa) with GPS/leveling ground control points. Int. J. Digit. Earth 12, 500–524. https://doi.org/10.1080/17538947.2018.1458163

Yu, H., Yang, Z., Tan, L., Wang, Y., Sun, W., Sun, M., Tang, Y., 2018. Methods and datasets on semantic segmentation: A review. Neurocomputing 304, 82–103. https://doi.org/10.1016/j.neucom.2018.03.037

Yue, L., Shen, H., Zhang, L., Zheng, X., Zhang, F., Yuan, Q., 2017. High-quality seamless DEM generation blending SRTM-1, ASTER GDEM v2 and ICESat/GLAS observations. ISPRS J. Photogramm. Remote Sens. 123, 20–34. https://doi.org/10.1016/j.isprsjprs.2016.11.002

Zhao, F., Xiong, L., Wang, C., Wang, H., Wei, H., Tang, G., 2021. Terraces mapping by using deep learning approach from remote sensing images and digital elevation models. Trans. GIS 25, 2438–2454. https://doi.org/10.1111/tgis.12824

Zhou, Z., Rahman Siddiquee, M.M., Tajbakhsh, N., Liang, J., 2018. UNet++: A Nested U-Net Architecture for Medical Image Segmentation, in: Stoyanov, D., Taylor, Z., Carneiro, G., Syeda-Mahmood, T., Martel, A., Maier-Hein, L., Tavares, J.M.R.S., Bradley, A., Papa, J.P., Belagiannis, V., Nascimento, J.C., Lu, Z., Conjeti, S., Moradi, M., Greenspan, H., Madabhushi, A. (Eds.), Miccai, Lecture Notes in Computer Science. Springer International Publishing, Cham, pp. 3–11. https://doi.org/10.1007/978-3-030-00889-5_1

Zhu, J., Taylor, T., Currens, J., Crawford, M., 2014. Improved Karst Sinkhole Mapping in Kentucky Using Lidar Techniques: A Pilot Study in Floyds. J. Cave Karst Stud. 76, 207–216. https://doi.org/10.4311/2013ES0135

Zumpano, V., Pisano, L., Parise, M., 2019. An integrated framework to identify and analyze karst sinkholes. Geomorphology 332, 213–225. https://doi.org/10.1016/j.geomorph.2019.02.013