Deep learning with multisite data reveals the lasting effects of soil type, tillage and vegetation history on biopore genesis

Han, Eusun; Kirkegaard, John A.; White, Rosemary; Smith, Abraham George; Thorup-Kristensen, Kristian; Kautz, Timo; Athmann, Miriam

Published in:
Geoderma

DOI:
10.1016/j.geoderma.2022.116072

Publication date:
2022

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY

Citation for published version (APA):
Deep learning with multisite data reveals the lasting effects of soil type, tillage and vegetation history on biopore genesis

Eusun Han a,b,*, John A. Kirkegaard b, Rosemary White b, Abraham George Smith c, Kristian Thorup-Kristensen a, Timo Kautz d, Miriam Athmann e

a University of Copenhagen, Department of Plant and Environmental Sciences, Højbjerggårds Allé 13, 2630 Taastrup, Denmark
b CSIRO Agriculture and Food, PO Box 1700, Canberra, ACT 2601, Australia
c University of Copenhagen, Department of Computer Science, Universitetsparken 1, 2100 Copenhagen Ø, Denmark
d Humboldt University of Berlin, Albrechts Daniel Thaer-Institut für Agrar- und Gartenbauwiss, Unter den Linden 6, 10099 Berlin, Australia
e University of Kassel, Organic Farming and Cropping Systems, Nordbahnhofstraße 1a, 37213 Witzenhausen, Germany

A R T I C L E   I N F O

Handling Editor: Yvan Capowiez

Keyword:
Subsoil
Deep tillage
Convolutional neural network
Perennial crops
RootPainter
AI

A B S T R A C T

Soil biopore genesis is a dynamic and context-dependent process. Yet integrative investigations of biopore genesis under varying soil type, tillage and vegetation history are rare. Recent advances in Machine Learning (ML) made faster and more accurate image analysis possible. We validated a model trained on Convolutional Neural Network (CNN) using a multisite dataset from varying soil types (Luvisol, Cambisol and Kandosol), tillage (deep ploughing and without deep ploughing) and vegetation history (taprooted and fibrous-rooted crops) to automatically predict biopore formation. The model trained on the multisite dataset outperformed individually trained single-site models, especially when the dataset contained images with noise and/or fewer biopores. Our model successfully replicated previously established treatment effects but provided new insights at more detailed scales and for different pore-size classes. These insights demonstrated that effects of deep ploughing on soil biopores can persist for more than 50 years and are more pronounced on the Luvisol rather than the Cambisol soil type. The effects of perennial fodder crops with high biopore generating capacity were also shown to persist for at least a decade. Re-growing the same fodder crops or a mixture with grass had no further impact on biopore density but generated a shift in pore-size classes from large to smaller biopores. We suspect this is likely to have resulted from three possible scenarios; (1) newly created fine pores (1–4 mm); (2) blockage of large-sized pores by earthworm faeces; (3) decrease in pore diameter. In summary, by using a single robust model trained on the multisite dataset, we were able to generate new insights on pore-size distribution as affected by site, vegetation, and deep ploughing. We have demonstrated that Deep Learning-based image analysis can provide easier biopore quantification and can generate models that provide novel insights across different research settings consistently and accurately.

1. Introduction

Soil biopores are tubular-shaped voids created by plant root penetration and movement of soil burrowers such as earthworms (Kautz, 2015). They function as an easy pathway for root growth of crop plants into deeper soil layers (Huang et al., 2020; Nakamoto, 2000), which can enhance nutrient and water uptake from the subsoil (Gaiser et al., 2012; Han et al., 2017; Jakobsen and Dexter, 1988)

There is uncertainty on how soil biopores function in arable fields (Cresswell and Kirkegaard, 1995). Depending on the soil and vegetation types, researchers have drawn contrasting conclusions about biopore impacts. By referring to numerous empirical experiments under Australian soils such as Kandosols, Passiourea (1991) suspected that roots growing into large-sized biopores could be detrimental for plant growth due to clumping of the roots and the reduced soil-root contact, especially when the roots are trapped inside the pores. European studies tend to find more favourable effects possibly because the lower strength soils such as Luvisols allow plant roots to escape from soil pores into deeper soil layers (see review by Kautz, 2015). While individual studies have their own merits and specific conclusions, more comprehensive and
integrated comparisons across studies would be helpful to understand the value of the changes in subsoil porosity induced by biopores across different climate and soil conditions. This is particularly important when considering how different crop root traits (e.g. root depth, density) are more effective under resource-limited environments (e.g. drought, low nutrient availability), so that improved management of subsoil structure (e.g. cover cropping, crop sequence, strategic tillage) can be devised. Such GxExM (Genotype × Environment × Management) approaches are increasingly promoted for more rapid improvements in system productivity (Hunt et al., 2021; Kirkegaard, 2019).

A pre-requisite to establish more integrated research programs is to develop consistent analytical procedures using the image datasets from different research groups that may vary in scope and scale. To date, quantifying biopore density/volume in field conditions has involved visual observation by marking/counting pores on relatively larger excavated horizontal/vertical surfaces (e.g. 0.5 × 0.5 m; Han et al., 2015) or on soil core samples (e.g. 10 × 10 mm; White and Kirkegaard, 2010). Imaging the surface area for pore quantification is also common (Glab et al., 2013; Lamande et al., 2003; Wuest, 2001), and can be useful for more detailed analysis of pore diameters (Huang et al., 2020). However, such image analysis is not automated and requires cumbersome manual measurement with generic software such as the Bersholt IMAGE Measurement Software (64-bit).

In addition to the excavation process, visual counting of the marked overlays or photographic images is a time-consuming process, especially when the research scale is large. Further to that, manual biopore quantification can lead to ambiguous determination of pore size distribution. As a result, the manual quantification of soil pores is often restricted to easily definable biopore sizes such as greater than 2 mm (Huang et al., 2020) or size classes may not even be considered (Fueki et al., 2012; Luo et al., 2010). Yet biopore size distribution was found to have a wide range from smaller than 0.001 mm (Bodner et al., 2014) to larger than 16 mm (McCallum et al., 2004).

Biopore formation in arable soils has been known to vary depending on soil type, vegetation, and tillage history. Luvisols with high silt and clay contents and with bulk density of 1.45 g cm⁻³ are known to provide ideal conditions for stable biopore formation under intensive taproot-dominated vegetation (Han et al., 2015). In fact, lower bulk density (e.g. below 1.0 g cm⁻³ on Ultisols) was found to provide less stable conditions for biopore formation (Dörner et al., 2010). Whereas well-structured soils such as red Kandosols was considered to have less opportunity for biopore formation by vegetation and soil fauna due to high soil strength (greater than 4 MPa) and predominantly kaolinitic clay acting as a cementing agent (White and Kirkegaard, 2010). In many cases disruption of soil structure by deep tillage practices decreased biopore formation (Ehlers, 1975; Ehlers et al., 1983; Wuest, 2001). Across the soil types, evidence suggests that vegetation with larger root systems and distinctive taproot architecture can increase biopore formation (Athmann et al., 2013; McCallum et al., 2004).

There are contrasting effects of soil conditions and vegetation history on biopore formation and inconsistency in methods used to quantify them. Therefore, a consistent parameter or a functional model to define biopore formation under varying conditions and scales of research would be valuable. Recent advances in a Machine Learning (ML) approach with a simplified and user-friendly interface has made faster and more accurate image analysis possible, especially for those without in-depth program language skills (Han et al., 2021). The earlier validation of biopore formation based on a Convolutional Neural Network (CNN) confirmed that training a model i.e., learning from labelled examples, for accurate prediction of biopore from an image dataset derived from a single experiment was possible after only 2 h of training (Smith et al., 2022).

Based on this method development, two new research questions have evolved: (1) can the simultaneous quantification of biopores from different soil types, vegetation and tillage histories be achieved using with an AI-based image analysis method, and (2) can the generated model be used to derive new insights from existing images as well as unseen datasets to investigate biologically meaningful research questions? To answer these questions, we trained a model using an integrated multisite biopore dataset from four different sites with three soil types, vegetation history and climate. We validated the accuracy of the model against manual counting, then we used the model to extract new insights from the previous studies by extending the pore-size classes. We quantified the effect of deep ploughing on the biopore density in two soil types. The model was also used to quantify biopores from images from an unseen dataset from a new experiment to reveal the effect of vegetation management in an arable field.

2. Materials and methods

2.1. Site description

For model training purposes, we combined the image datasets from three sites in Germany and one site in Australia. Detailed description of the site and soil characteristics of the four study sites is presented in Table 1.

### 2.1.1. Luvisol in Meckenheim, Germany

The study site near Meckenheim, North-Rhine Westphalia, Germany (50° 37’ 9” N 6° 59’ 29” E) was established on a Haplic Luvisol (developed from loess) at campus Klein-Altenford of University of Bonn. Lucerne (Medicago sativa L.), chicory (Cichorium intybus L.) and tall fescue (Festuca arundinacea Schreb.) were grown for 1/2/3 years from 2009 to 2012 at sowing densities of 25, 5 and 30 kg ha⁻¹, respectively. The Festuca arundinacea was cut and mulched up to four times and terminated prior to biopore investigation in March 2012. See Table 2 for treatment overview in Meckenheim.

In 2018, lucerne (25 kg ha⁻¹), chicory (8 kg ha⁻¹) and tall fescue (30 kg ha⁻¹) were sown as sole crops on August 24, 2018 and grown for two consecutive years on each of the previous treatments established in 2012. Also, lucerne and tall fescue were sown as a mixture with the sowing densities of 21 kg ha⁻¹ and 9 kg ha⁻¹, respectively. Chicory (5 kg ha⁻¹) and tall fescue (15 kg ha⁻¹) were also sown as a mixture. Clover grass mixture (DSV Country Öko 2251) consisting 30 % of *Lotus perenniv*, 20 % of *Lotus multiflorum*, 20 % of *Lotus hybridum* and 30 % of *Trifolium pratense* L. was sown with a sowing density of 35 kg ha⁻¹. All fodder crops and mixtures were mulched three times in 2018 and 2019 and tilled in March 2020 after biopore counting.

For this study, we disregarded the effect of crop duration from the Meckenheim dataset, and treated the 1 year, 2 years and 3 years-treatments equally (See Table 2). This was based on the original study, which found no effect on biopore density after a longer crop duration of fodder crops over a shorter crop duration (Han et al., 2015).

### 2.1.2. Luvisol in Banteln, Germany

The study site in Banteln, Lower Saxony in Germany (52° 13’ 13”N 9° 44’ 56”E) was established on a Haplic Luvisol. Inverse deep ploughing was carried out in 1965 and control plots without deep plough remained in parallel. Since then, the site has been managed as a cropland grown with sugar beet, winter wheat and mustard.

### 2.1.3. Cambisol in Elze, Germany

The study site in Elze, Lower Saxony in Germany (52° 35’ 06”N 9° 45’ 29”E) was established on a Dystric Cambisol. Similar to the study site in Banteln, the deep ploughing was performed in 1968. Various crops such as winter canola, winter rye, potato and mustard have been grown since then.

### 2.1.4. Kandosol in Bethungra, NSW Australia

The study site near Bethungra, NSW in Australia (34° 43’5S, 147° 48’E) was established on a red Kandosol. Prior to biopore investigation, lucerne was grown as a pasture in 1998–2003, after which wheat...
Table 1
Soil characteristics and agronomic activities.

<table>
<thead>
<tr>
<th>Soil type</th>
<th>Site</th>
<th>Depths of investigation</th>
<th>Agronomic activity</th>
<th>Sand (%)</th>
<th>Silt (%)</th>
<th>Clay (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haplic Luvisol*</td>
<td>Meckenheim, Germany</td>
<td>45 cm</td>
<td>Fodder cropping with lucerne, chicory, tall fescue and red clover for 1/2/3 years</td>
<td>7</td>
<td>75</td>
<td>17</td>
</tr>
<tr>
<td>Haplic Luvisol**</td>
<td>Banteln, Germany</td>
<td>40, 60, 80 cm</td>
<td>Inversive deep ploughing (85 cm depth) in 1965</td>
<td>5</td>
<td>82</td>
<td>13</td>
</tr>
<tr>
<td>Dystric Cambisol**</td>
<td>Elze, Germany</td>
<td>40, 60, 80 cm</td>
<td>Inversive deep ploughing (55 cm depth) in 1968</td>
<td>84</td>
<td>12</td>
<td>4</td>
</tr>
<tr>
<td>Red Kandosol***</td>
<td>Bethungra, Australia</td>
<td>10–40, 50–80, 90–120, 130–160 cm</td>
<td>Lucerne-pasture for 6 years</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* From Vetterlein (2013).
** From Alcantara et al. (2017).
*** From Kirkegaard et al. (2007).

Table 2
Treatment overview in Meckenheim, Germany.

<table>
<thead>
<tr>
<th>Agronomic activity</th>
<th>Fodder cropping 2009–2011*</th>
<th>Fodder cropping 2019–2020</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lucerne</td>
<td>Lucerne grown for 1 year</td>
<td>Lucerne + Tall fescue grown for 2 years</td>
</tr>
<tr>
<td></td>
<td>Lucerne grown for 2 years</td>
<td>Lucerne re-grown for 2 years</td>
</tr>
<tr>
<td></td>
<td>Lucerne grown for 3 years</td>
<td></td>
</tr>
<tr>
<td>Chicory</td>
<td>Chicory grown for 1 year</td>
<td>Chicory + Tall fescue grown for 2 years</td>
</tr>
<tr>
<td></td>
<td>Chicory grown for 2 years</td>
<td>Chicory re-grown for 2 years</td>
</tr>
<tr>
<td></td>
<td>Chicory grown for 3 years</td>
<td></td>
</tr>
<tr>
<td>Tall fescue</td>
<td>Tall fescue grown for 1 year</td>
<td>Clover + Grass grown for 2 years</td>
</tr>
<tr>
<td></td>
<td>Tall fescue grown for 2 years</td>
<td>Tall fescue grown for 2 years</td>
</tr>
<tr>
<td></td>
<td>Tall fescue grown for 3 years</td>
<td></td>
</tr>
</tbody>
</table>

*For this study, we disregarded the crop duration for analysis as the original study did not show effects.


2.2. Acquisition of biopore images

2.2.1. German sites (Meckenheim, Banteln, Elze)

In 2012, the images were collected from all the 9 treatments with 12 replicates totalling 108 images (see Table 2 and3). The imaging area on Luvisol in Meckenheim was prepared by excavating a horizontal surface to the depth of 45 cm. The exposed surface area of 0.25 m² (50 × 50 cm) was photographed with Panasonic DMC-TZ10 in 2012 (2141 × 2141 pixels: 180 DPI) in 2012, and with NIKON D7100 in 2020 (3660 × 3660 pixels: 300 DPI) in 2020. Only the plots that were re-grown with the fodder crops or mixture with clover and grass were investigated in 2020 having 6 treatments with 4 replicates plus 4 control images (28 in total). The plots and depth-points were located by pre-recorded GPS coordinates and the buried marking pins at four corners at 45 cm. Similar image acquisition was done on Luvisol in Banteln (0.2–1 m) and Cambisol in Elze (0.2–0.8 m) in 2019 at each 20 cm interval with 3 replicates, totalling 30 and 24 images, respectively (Table 3). Except for the images taken in 2020 from Luvisol in Meckenheim, all other images were manually counted for biopore numbers. Pore-size classes larger than 2 mm were quantified during the manual measurement.

2.2.2. Australian site (Bethungra)

Biopore images from Kandosol in Bethungra were acquired from cross-sections of soil cores in 2005 from 10 to 160 cm depth using thin-walled steel tubes of 9.4 cm in diameter and 40 cm in length. There were totally 14 cores collected from which the cross-sections at 1 or 2 depth-levels were investigated totalling 239 images (Table 3). Image capture was carried out using a Zeiss Axiocam digital camera with Axiosvision software (Carl Zeiss Australia, Sydney, Australia). Ten random images of 1 cm² in size from one cross-section were captured at a resolution of 2600 × 2060 pixels, which resulted in images not in focus with overlapped areas. The images were combined into a single image with most regions of the image in focus using AutoMontage Essentials (Syncroscopy, Frederick, MD, USA). Detailed description on image acquisition can be found in White and Kirkegaard (2010). The manual counts were done on core surface and provided by the authors, who categorized pore-size classes: <0.2 mm, 0.2–1.0 mm and greater than 1.0 mm.

2.3. Software and training strategy

We used the RootPainter software, an image analysis software using CNN based on U-Net (Ronneberger et al., 2015). The software features a user-friendly GUI (Graphical User Interface) that allows users to interact with on-going training process, which allows the users input to be used for inference called corrective annotation. Detailed description of the software is available in Smith et al. (2022).

2.4. Preparation of training dataset

Training datasets were prepared from the original images in JPEG format by using the “Create training dataset” function in the software (Table 3). For efficient training, the original images from German sites were tiled into four separate images with 1000 pixels in width. Fifteen out of 108 images from Luvisol 2012 in Meckenheim were captured in low resolution (804 × 804 pixels) which were not tiled. The Australian dataset was not tiled as the images were captured at a small scale, and their pixel size was reduced to 900 instead for efficient annotation/training.

Table 3
Training dataset description.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Imaging</th>
<th>Image numbers</th>
<th>Tile size (pixels)</th>
<th>Tile number</th>
<th>Format</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Imaging</td>
<td>Original image</td>
<td>Single-site dataset</td>
<td>Multisite dataset</td>
<td></td>
</tr>
<tr>
<td>Luvisol 2012 in Meckenheim</td>
<td>Excavated horizontal surface</td>
<td>108</td>
<td>387</td>
<td>96</td>
<td>1000</td>
</tr>
<tr>
<td>Luvisol 2019 in Banteln</td>
<td>Excavated horizontal surface</td>
<td>30</td>
<td>120</td>
<td>96</td>
<td>1000</td>
</tr>
<tr>
<td>Cambisol 2019 in Elze</td>
<td>Excavated horizontal surface</td>
<td>24</td>
<td>96</td>
<td>96</td>
<td>1000</td>
</tr>
<tr>
<td>Kandosol 2005 in Bethungra</td>
<td>Cross-sections of soil cores</td>
<td>239</td>
<td>239</td>
<td>107</td>
<td>900</td>
</tr>
</tbody>
</table>
2.5. Training

Single-site models: Datasets from Luvisol 2012 in Meckenheim, Luvisol 2019 in Banteln, Cambisol 2019 in Elze and Kandosol 2005 in Bethungra were used individually for model training with 2 h annotation time – the time spent for marking foreground and background using RootPainter interface.

Multisite model: The four training datasets were combined. The minimum number of images to be shared was set to 96, which was the maximum number of images from from Cambisol 2019 in Elze. This resulted in 96 M pixels to be shared by each dataset. The training dataset from Kandosol 2005 in Bethungra contained only 900 × 900 pixels, thus, 107 images were shared (96.3 M pixels in total). The low-resolution images from Luvisol 2012 in Meckenheim were not included in the combined dataset. The annotation time was restricted to 2 h, after which training continued until a new model was generated.

2.6. Data extraction

The original datasets were segmented using the generated models from each training run. The segmentation was used to extract “region properties” in csv format containing the area of each connected region in pixels. The area was used to further classify the biopsy size classes using the calibration factor from the image acquisition, after which the number of segmented biopores per image was computed.

2.7. Model validation

2.7.1. Validation I: Correlation with manual counts

The manual counts were used for linear regression analysis against the segmentation generating R² values for each model, which were used to evaluate model accuracy. For this purpose, we used the number of biopores larger than 2 mm for the German datasets due to the ambiguity to define the pores smaller than 2 mm. For the Australian dataset, as the original counting was done, only the number of biopores larger than 0.2 mm was used for validation.

2.7.2. Validation II: Application of model to the trained datasets

Extended analysis: The original studies did not compare the density of pores smaller than 2 mm (Han et al., 2015) and 0.2 mm (White and Kirkegaard, 2010) due to the ambiguity in defining them on the images obtained. Using the multisite model, we re-analyzed the legacy datasets at more detailed size-classes. Specifically, biopsy size classes from 1 to 10 mm was analyzed for Meckenheim 2012 dataset which is an extension from 1 to 5 mm in Han et al. (2015). From Kandosol 2005 datasets, we were able to extract the number of pores smaller than 0.1 mm and larger than 0.5 mm, which were not analyzed previously. For 0.1 mm size-class, we have applied a threshold to filter the segmentation smaller than 0.01 mm.

Treatment effects: Datasets were analyzed using the multisite model to reveal the effects of (1) fodder cropping (Luvisol 2021 in Meckenheim); (2) soil depth (Kandosol 2005 in Bethungra); (3) tillage (Luvisol 2019 in Banteln) on biopsy density.

Site comparisons: Using the multisite model, we compared the biopsy density of size-classes from 1.0 mm to 10.0 mm in 0.5 mm increments between the four sites from a similar depth in the soil profile (40–50 cm).

2.7.3. Validation III: Application of model to an unseen dataset

We used the multisite model to generate new results from an unseen dataset - meaning the images were not used in training. The newly generated biopsy density from Luvisol 2020 in Meckenheim, was compared between the treatments. The previous treatments in 2012 were growing lucerne, chicory and tall fescue mixes for either 1, 2 or 3 years initiated in 2009. From 2018, the same sole crops (lucerne, chicory and tall fescue) were grown for two years on the original treatments (see Table 2). Additionally crop mixtures of lucerne and tall fescue, chicory and tall fescue were implemented on lucerne 2012 and chicory 2012 plots. In tall fescue 2012 treatments an additional red clover treatment was implemented.

2.8. Statistics

R version 4.0.1 (R Core Development Team, 2019) was used for statistical analysis. The segmentation from all the models from each training run was used for linear regression analysis (lm function) against the manual counts generating R² values for each model, which were used to evaluate performance. Significance of effects (fodder cropping, deep ploughing, study sites) were measured by a mixed-effects model based on the approximated degrees of freedom calculated by the package lmerTest followed by multiple comparisons (multcomp package) (P ≤ 0.05). Replicates were assigned as random variables. Exponential, quadratic, and cubic equations (lm function) were used to fit pore numbers as a function of soil depth for the dataset from Kandosol in Bethungra.

3. Results

3.1. Validation I: Correlation with manual counts

We report the segmentation model performance measured by regression analysis between the manual counts vs segmentation on each dataset for each single-site model and the multisite model. We observed improved accuracy with the multisite model compared with the single-site model, leading to greater R² with the same length of training/annotation time (Fig. 1a-c). We have visually observed that the increase in performance was due to the better prediction of biopores from datasets containing topsoil layers, which often resulted in false positives on non-biopore objects – e.g. Luvisol in Banteln (Fig. 1d vs 1g; R² from 0.8 to 0.9). The multisite model had the highest accuracy for all the datasets with lower biopsy density and background noise such as from the Cambisol in Elze (Fig. 1e vs 1h; R² from 0.3 to 0.6) and Kandosol in Bethungra (Fig. 1f vs 1i; R² from 0.1 to 0.6). The dataset from Luvisol in Meckenheim with distinctive biopores had an R² of 0.7 for the single-site model and did not benefit from analysis with the multisite model (R² = 0.6).

3.2. Validation II: Application of model to the trained datasets

3.2.1. Crop species

We analyzed the effects of cultivating the three fodder crops from the original dataset (Han et al., 2015) with more detailed pore-size classes using the multisite model. In the original study, based on visual observations, we concluded that growing the taprooted crops (lucerne and chicory) led to a significant increase in biopsy density compared to the fibrous-rooted tall fescue.

From our analysis using automated segmentation, the effects of fodder cropping on biopsy density were restricted to pores in the 2.5 mm to 5.0 mm range (Fig. 2). While lucerne and tall fescue remained - the strongest and the weakest biopsy-generating species, respectively, the results with chicory were rather variable. As a result, the total biopsy density (1.0–10.0 mm) showed the effects of fodder cropping, in which lucerne cultivation led to greater biopsy density followed by chicory and tall fescue cultivation (Fig. 2).

3.2.2. Soil depth

In the original study (White and Kirkegaard, 2010), the authors concluded that the number of pores of all size classes (<0.2 mm, 0.2–1.0 mm and greater than 1.0 mm) decreased with depth. It was highlighted that the smaller pores (<0.2 mm) decreased more rapidly than the larger pores (0.2–1.0 mm). We determined the same parameter using the multisite model, which allowed us to classify pores into more detailed
categories, namely, <0.1 mm, 0.1–0.2 mm, 0.2–0.3 mm, 0.3–0.4 mm, 0.4–0.5 mm and greater than 0.5 mm. We have found similar trends after fitting the three functions to the data (exponential, quadratic and cubic), where pore numbers decreased with depth, but the trends differed between the pore-size classes (Fig. 3). Smaller pores, namely, <0.1 mm, 0.1–0.2 mm and 0.2–0.3 mm showed higher $R^2$ values when plotted against soil depth - ranging from 0.7 to 0.8. $R^2$ values for the larger pores ranged from 0.2 to 0.4. In general, the cubic function predicted the relationship better than exponential or quadratic fits.

One of the new insights was that the pores smaller than 0.1 mm were as abundant as 0.1–0.2 mm and 0.2–0.3 mm-sized pores – which could not be confirmed in the previous studies. Except for 50–80 cm depth, pore size of <0.1 mm was the most dominant pore size classes at all depths – comprising on average 26 % of all pores across the soil depth.

3.2.3. Tillage

We used the multisite model to segment the datasets from Luvisol 2019 Banteln and Cambisol 2019 Elze to compare the effects of the inversive deep ploughing on biopore density at five depths. Surprisingly, very few biopores were found from Cambisol in the Elze dataset for...
either treatment (data not shown). Only 16 biopores per m\(^2\) ranging from 2.5 to 5.0 mm at 0.4 m soil depth were detected from treatments with deep ploughing. For plots without deep ploughing, 4 (2.0–2.5 mm) and 16 biopores (1.0–3.0 mm) at 0.4 and 0.6 m were found, respectively.

From the Luvisol 2019 Banteln dataset, we compared the biopore density with and without deep ploughing at five depth-levels for 18 biopore-size classes (1.0 to 10.0 mm). The most profound differences between the treatments were observed at 0.6 m (Fig. 4). Effects of deep ploughing were significant for biopore size classes of 2.0–2.5 mm, 2.5–3.0 mm, 3.5–4.0 mm, 4.5–5.0 mm and 5.0–5.5 mm, in which treatments without deep ploughing had a greater biopore density. Some effects were shown at 0.4 m (1.0–1.5 mm and 5.0–5.5 mm), at 0.8 m (7.0–7.5 mm and 8.0–8.5 mm) and at 1.0 m depth (7.5–8.0 mm). We noticed that in the subsoil layers biopores larger than 8.5 mm were observed mainly for treatments without deep ploughing. The total biopore density (1.0–10.0 mm) was affected by deep ploughing at 0.8 m.

### 3.2.4. Soil types and biopores

Using the multisite model, we were able to make a detailed comparison of pore-size distribution from 1.0 mm to 10.0 mm biopore sizes at each 0.5 m intervals between the four sites based on a single model. We excluded the pore-size classes from 0.1 mm to 1.0 mm as other than the datasets from Bethungra, the mean pixel number per mm ranged from 4 to 7. In general, Luvisols from two sites had a dominant density of biopores with sizes larger than 1.0 mm dominated compared with Cambisol and Kandosol (Fig. 5). The Kandosol exhibited an extremely high biopore density smaller than 1.0 mm, and a low density in size-classes larger than 1.0 mm. The Cambisol contained a small number of pores compared to other soil types across all biopore sizes. Between the two Luvisols, the soil treated with long-term fodder cropping in

![Fig. 3. Re-analyzed biopore density from more pore size classes to depth using the dataset from Kandosol 2005 in Bethungra (White and Kirkegaard, 2010). Exponential (left), quadratic (middle) and cubic equations (right) were fitted between biopore density (<0.1, 0.1–0.2 mm, 0.2–0.3 mm, 0.3–0.4 mm, 0.4–0.5 mm and greater than 0.5 mm) and soil depth (m).](image-url)
Meckenheim tended to have a greater number of large-sized biopores (8.5–9.0 mm and 9.5–10.0 mm) compared to the soil in Banteln, however, the differences were not significant. The total biopore density (1.0 mm to 10.0 mm) did not differ between the two Luvisols and Kandosol. The Cambisol exhibited the fewest pores, which was significantly different from the two Luvisols.

Fig. 4. Biopore density (number of biopores per m$^2$) as affected by tillage treatments for the Luvisol at Banteln in 2019. Asterisks indicate significant effects of tillage within biopore size class ($P \leq 0.05^*; P \leq 0.01^{**}; P \leq 0.001^{***}$; Mixed-effects models).

Fig. 5. Biopore-size distribution (1.0–10.0 mm) in Luvisol 2012 in Meckenheim (45 cm), Luvisol 2019 in Banteln (40 cm), Cambisol 2019 in Elze (40 cm), Kandosol 2005 in Bethungra (averaged 30, 31, 57 and 59 cm). Small letters indicate significant differences between the soil type within the biopore size ($P \leq 0.05$; Tukey HSD).
3.3. Validation III: Application of model to the unseen dataset

Using the multisite model, we segmented the unseen dataset, which was not used for training. The dataset from Luvisol in 2020 was collected from the same site in Meckenheim, where sole-crops and crop mixtures were grown on top of the previously implemented treatments with lucerne, chicory and tall fescue in 2012. There was an increase in biopore density of 1–2 mm size when lucerne and mixture of lucerne and tall fescue were grown on the plots where lucerne was grown in 2012, whereas no further effects were shown in other biopore sizes (Fig. 6).

Re-growing chicory and mixture of chicory and tall fescue led to an increase in biopore density of 1–2 mm in size (Fig. 6). Re-growing chicory also increased biopore density of 2–5 mm, whereas growing chicory and tall fescue as a mixture resulted in a moderate increase only.

Growing clover and grass as a mixture significantly increased the number of biopores of 1–2 mm size in plots where tall fescue was cultivated (Fig. 6). A significant decrease in large-sized biopores (greater than 5 mm) was noted after tall fescue as a sole-crop was re-grown. The total biopore density (1.0 mm-10.0 mm) remained unchanged for all the treatments, whereas the effects of fodder cropping implemented in 2012 were still seen in 2020. Chicory 2012 treatments (1.0 mm-10.0 mm) exhibited a greater biopore density compared with tall fescue 2012 when measured after implementation of the treatments in 2020. Lucerne grown in 2012 showed an intermediate effect in 2020.

4. Discussion

4.1. Model performance

We demonstrated faster and more accurate quantification of soil biopore density using the multisite model that was trained on a combined dataset comprising of different soil types, vegetation history and scales of sampling. The approach was made possible using an interactive-ML approach for segmentation. Training on the multisite dataset did not only save time by reducing the necessity to train for each dataset (1 interactive training session for 2 h vs 4 training sessions totalling 8 h), but it also improved model performance, especially for the challenging datasets with limited biopore numbers.

We believe that the improved performance can be explained by the principle of multitask learning due to the amounts of variation present in our multisite dataset. Multitask training is known to be an effective form of training that can introduce an inductive bias which can make the model prefer hypotheses that explain more than one task (Ruder, 2017).

From our examples, it is not possible to precisely locate the source of the observed improvement in model performance. Nevertheless, visual observation during training confirmed the following. Firstly, it seems that the clear examples from two Luvisols may have helped the model to learn the “roundness” of biopores. This roundness-bias or learned biopore shape prior likely enabled improved performance on the Kandosol.
2005 in the Bethungra dataset, where cracks and fissures were abundant, obscuring the prototypical biopore shape. When separately trained, the model from the Cambisol 2019 in Elze struggled due to the false negatives on the fine pores, especially on treatments with deep ploughing having mixed soil colour. We believe that both the clear examples of biopores from the two Luvisols, as well as the fine pores depicted on varying soil colour from the Kandosol in the Bethungra dataset contributed to the improvement when analyzing the data with the multisite model. Finally, the performance of the model for the dataset Luvisol 2019 in Banteln also improved when using a multisite dataset. We assume that the model learned about pore segmentation on topsoil images from other datasets containing extremely fine pores such as Kandosol 2005 in Bethungra dataset.

We also suspect that the interactive-ML approach contributed to the efficiency of the multisite annotation process. As corrective-annotation was used, the labels were allocated between the constituent datasets based on the magnitude of the model error throughout training, enabling targeted annotation towards under-represented features in the combined data.

Our primary goal from merging the datasets was not to improve performance, but to complete the same tasks, i.e., detection of biopores, more efficiently. Usually biopore quantification in field scales requires several counters, and more so for the multisite investigations. This can potentially cause inter-observer variation. Using RootPainter, we overcame this by reducing the observer to a single person which then can provide consistent perspective on biopore quantification. Nevertheless, our observation suggests that training on multisite dataset with variable images improved the biopore segmentation model generalization performance. This in turn improved biopore predictions on the more challenging datasets. Our observations align with previous studies on medical data (Amyar et al., 2020).

### 4.2. New insights from the legacy studies

Using the segmentation, we confirmed the treatment effects identified in the original studies (Han et al., 2015; White and Kirkegaard, 2010), in which altered biopore density as a function of crop root systems and soil depth was demonstrated. However, we were also able to generate new results beyond the scale and scope of the original research. For example, in Han et al. (2015), biopores smaller than 2 mm were not investigated. Moreover, pores smaller than 0.1 mm could not be quantified in the original study at Bethungra (White and Kirkegaard, 2010) - limited by the time and cost of manual observation in both cases. In the case of the Han et al. (2015) study, our segmentation on the dataset revealed that effects of fodder cropping were rather restricted mainly to a specific range of pore-sizes (greater than 2.5 to < 5.0 mm). The effect on these large-sized pores has been observed in previous studies (McCallum et al., 2004; Perkins et al., 2014). We also expected to see a reversed effect on smaller pores (< 2 mm), i.e. increased density of small-sized pores when fibrous rooted tall fescue was cultivated – but this was not shown from our segmentation. We assume that this was caused by the similar number/density of fine roots responsible for making such small pores between the fibrous and taproot systems (Han et al., 2017). Also considering the high proportion of this root diameter class for both root systems (greater than 97 %), there is still a possibility that the effects might go beyond the scale of this study, i.e. < 1 mm.

For the Australian Kandosol the results support our assumption on the relationship between vegetation and fine pores. We found that the pore-size below 0.1 mm is one of the dominant size classes and exhibited a similar change with to soil depth as the two other pore-size classes (0.1–0.2 mm and 0.2–0.3 mm). According to previous studies on loamy soils some of the very fine pores (<0.005 mm) were affected by both plant root systems (Bodner et al., 2014) and earthworm activities (Lamande et al., 2003). Usually, biopore density tends to decrease with soil depth (Perkins et al., 2014), which is why we believe that such fine pores shown in Kandosol 2005 should be regarded as biopores created by vegetation.

#### 4.3. Deep tillage effect persisting for 50 years

Our results indicate a long-term effect of invasive deep ploughing on subsoil structure. Tillage at depth is known to discourage the formation of soil biopores (Ehlers et al., 1983), however, this might be the first report demonstrating that the effects of deep ploughing on soil porosity can persist for over half a century after implementation.

Deep tillage practices have been tested under a range of soil types to alleviate limited root growth in compacted layers (Hartmann et al., 2008). Deep ploughing effectively reduces soil strength, which can
promote root growth and improve crop productivity by allowing plants to access subsoil resources (Cai et al., 2014). However, effects of deep tillage on crop yield are context-dependent. A recent review by Schneider et al. (2017) performed a meta-analysis in this regard involving 67 experimental sites in temperate latitudes. They found that crop yield response to deep tillage was site-specific. The results demonstrated that deep tillage practices had highest potential for increasing crop yield when the sites suffered from compacted soil layers and dry conditions during crop growth. Sites without these conditions exhibited inconsistent effects. The risk of having negative effects after deep tillage became more pronounced in soils with greater than 70 % growth of crop plants. This generates some uncertainty regarding the specific addressed on how soil types affect biopore density, and root density (D. López et al., 2014; Kautz, 2015). Soil type and other factors lead to variation in composition of the soil and tillage practices (Alouaïri et al., 2011; Bodner et al., 2014; Kautz, 2015). Soil type and other factors lead to variation in bulk density, and soils with high bulk density retain pores of larger size with higher continuity, whereas fine pores dominate under low bulk density (Dörner et al., 2010). However, previous studies have not specifically addressed on how soil types affect biopore density, and root growth of crop plants. This generates some uncertainty regarding the legacy effects on subsoil structure in different studies done on different soil types.

We used a single multisite model to compare soil biopore density of 4 different sites. Two extreme cases were the Kandosol and Cambisol. The majority of pores in Kandosol belonged to the 1.0–2.0 mm size-class, whereas Cambisol exhibited very few pores across all biopore size classes. The Kandosol at the study site is known to have a well-structured subsoil prone to chemical cementing (White and Kirkegaard, 2010), which might have affected the stability of larger-sized pores, or the plants might have preferably grown roots into the pores, fissures and cracks as shown in the original study.

In contrast, no high soil strength has been noted for the Cambisol. However, the soil type consisted of a high proportion of sand (84 %) and had a low share of clay (4 %), which might have caused less stability to initially formed biopores (Dörner et al., 2010).

The two Luvisols showed a similar gradient of biopore density across biopore size classes. Both Luvisols derived from loess with high clayey content which are known to favour biopore formation (Kautz et al., 2010; Perkins et al., 2014).

4.4. Different soil types, different biopore density

Soil type is assumed to be a main driver defining pore stability and continuity, in combination with climatic conditions (wetting and drying), vegetation (root growth and earthworm activity), chemical composition of the soil and tillage practices (Alouaïri et al., 2011; Bodner et al., 2014; Kautz, 2015). Soil type and other factors lead to variation in bulk density, and soils with high bulk density retain pores of larger size with higher continuity, whereas fine pores dominate under low bulk density (Dörner et al., 2010). However, previous studies have not specifically addressed on how soil types affect biopore density, and root growth of crop plants. This generates some uncertainty regarding the legacy effects on subsoil structure in different studies done on different soil types.

We used a single multisite model to compare soil biopore density of 4 different sites. Two extreme cases were the Kandosol and Cambisol. The majority of pores in Kandosol belonged to the 1.0–2.0 mm size-class, whereas Cambisol exhibited very few pores across all biopore size classes. The Kandosol at the study site is known to have a well-structured subsoil prone to chemical cementing (White and Kirkegaard, 2010), which might have affected the stability of larger-sized pores, or the plants might have preferably grown roots into the pores, fissures and cracks as shown in the original study.

In contrast, no high soil strength has been noted for the Cambisol. However, the soil type consisted of a high proportion of sand (84 %) and had a low share of clay (4 %), which might have caused less stability to initially formed biopores (Dörner et al., 2010).

The two Luvisols showed a similar gradient of biopore density across biopore size classes. Both Luvisols derived from loess with high clayey content which are known to favour biopore formation (Kautz et al., 2010; Perkins et al., 2014).

4.5. Re-bioporing – Is it size-shifting rather than creation?

The multisite model was used to segment biopores on the unseen dataset from Luvisol 2020 in Meckenheim. The experiment in 2020 aimed to measure the biopore density as affected by re-growing the fodder crops in different mixtures on the same site used in 2012 (Han et al., 2015). It has been reported that the share of crop roots inside biopores can range from 20 % to 47 % (e.g. wheat and barley) in case of loamy soil (Han et al. 2015; Perkins et al. 2014; Han et al. 2017). Given that the fodder crops will also re-utilize pore channels for root growth, our main question was if new biopores can be formed using the same cover crop and cover crop mixture under the pre-existing biopore-rich condition. Our results indicated that the total biopore density (1–10 mm) remained un-changed after the re-cultivation. However, we found that the re-cultivation of fodder crops increased the numbers of smaller pores roughly from 1.0 to 3.5 mm, meanwhile the pore numbers of the large size classes (7.0–9.5 mm) had declined after the re-cultivation in 2020.

Based on visual observation, we identified three possible scenarios that might explain these results. Firstly, new pores of relatively smaller sizes (1.0 to 4.0 mm) did appear in datasets acquired in 2020 (Fig. 7a), which would have partially contributed to the increase in pore numbers in those size-classes. Secondly, we found that some large pores became blocked by earthworm faeces (Fig. 7b) or disappeared completely between the two observation points (Fig. 7c). Finally, we observed decreases in pore diameter from 2012 to 2020 in most of the cases we measured (Fig. 7a-c). From the results and observations, we are unable to conclude that the changes in biopore density in the different size-classes are due to the creation of new biopores – as the total biopore density remained un-changed. Moreover, all the three scenarios were observed regardless of the treatment, which defeats the effects of contrasting root systems.

Instead, we conclude that this shift of biopore size-classes over time from larger sizes to smaller sizes as caused by other factors, such as root decay creating biopores which can progress over one to several years depending on the root morphology, crop species and N regime (Herrera et al., 2017; Zhang and Wang, 2015), or growing of annual crops between 2012 and 2018 (e.g. wheat, barley). The deposition of earthworm casts may also have affected the dynamics of pore formation (Pagenkemper et al., 2015) as we observed.

In 2020, biopore quantification was undertaken while the biopore-generating fodder crops were still grown on the field. Considering possible total or partial blocking of biopores by fodder crop roots or earthworm cast, it may be reasonable to quantify biopores again on the same sites after decay of fodder crop roots and after a period of regular plowing where promotion of anecic earthworms has ceased.

Surprisingly, the treatment effects created in 2012 were still dominating in 2020, especially for chicory plots, which revealed a greater number of total biopores compared with the tall fescue plots. This implies that perennial fodder cropping can have a stable and long-term effect on biopore formation in arable subsoil. Such a long-term effect can be beneficial as annual crops in several crop rotation cycles can grow deeper roots for better subsoil resource acquisition (Han et al., 2017).

5. Conclusions

We performed a three-step validation using the CNN-based image analysis procedure for faster and more accurate biopore quantification. Firstly, training on datasets combined from four different experiments accurately quantified the biopores more rapidly than using each dataset for training separately. Secondly, the multisite model revealed similar treatment effects to those originally observed but revealed new insights when applied to two legacy datasets. Finally, biologically meaningful results were also obtained when the model was applied to an unseen dataset from more recent experiment that was not included in training process.

Our results suggest that the segmentation made more extensive and detailed measurement of biopore size classes possible and revealed several interesting new insights. These included the observations that taprooted crops influenced formation of biopores < 2 mm in the subsoil. Such a trend was also seen when the crops were re-grown in areas in which they had produced the biopores previously. In summary, using a single robust model trained on an integrative dataset, we were able to compare pore-size distribution between the different sites with differing vegetation, and revealed long-term effects of deep tillage. We conclude that CNN-based image analysis can lead to easier biopore quantification across research platforms and experiments by capturing the necessary variation in the datasets used for training.

Author contributions

EH and MA co-designed the research. EH conducted the research and...
prepared the manuscript. JK and RW provided guidance for recognition of the Australian biopores and performed a quality check on segmentation. TK and MA coordinated the manual counting of biopores from German sites. AGS provided technical advice on Machine Learning. The GPU server was provided by KTK. All authors contributed to manuscript writing.

Funding

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 884346, SenseFuture. The biopore images from Meckenheim in 2012 and 2020 were obtained under the FOR 1320 project funded by German Research Foundation (DFG) and BonaRes research project Soil3 (Grant No 031B0515E) funded by the German Federal Ministry of Education and Research (BMBF), respectively. Use of GPU server for Machine Learning was supported by the project DeepFrontier (funded by Villum Foundation; grant number VRK023338).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Biopore images, manual counts, training datasets, and models generated after training are available at https://doi.org/10.5281/zenodo.6952038.

References

