Computational and Structural Biotechnology Journal 21 (2023) 3696–3704

Contents lists available at ScienceDirect

Computational and Structural Biotechnology Journal

journal homepage: www.elsevier.com/locate/csbj

Research article

IMFSegNet: Cost-effective and objective quantification of intramuscular fat in histological sections by deep learning

Jan-Philipp Praetorius [a,b], Kassandra Walluks [a,b,c], Carl-Magnus Svensson [a], Dirk Arnold [d,e], Marc Thilo Figge [a,f,*]

[a] Applied Systems Biology, Leibniz institute for natural Product Research and infection Biology – Hans Knöll institute (HKI), Jena, Germany
[b] Faculty of Biological Sciences, Friedrich Schiller University Jena, Jena, Germany
[c] Institute of Zoology and Evolutionary Research, Faculty of Biological Sciences, Friedrich Schiller University Jena, Jena, Germany
[d] Department of Otorhinolaryngology, Jena University Hospital, Jena, Germany
[e] Facial-Nerve-Center Jena, Jena University Hospital, Jena, Germany
[f] Institute of Microbiology, Faculty of Biological Sciences, Friedrich Schiller University Jena, Jena, Germany

ARTICLE INFO

Key words:
Biomedical image analysis
Intramuscular fat
Deep Learning
Explainable artificial intelligence
Muscle biopsies
H&E staining

ABSTRACT

The assessment of muscle condition is of great importance in various research areas. In particular, evaluating the degree of intramuscular fat (IMF) in tissue sections is a challenging task, which today is still mostly performed qualitatively or quantitatively by a highly subjective and error-prone manual analysis. We here realize the mission to make automated IMF analysis possible that (i) minimizes subjectivity, (ii) provides accurate and quantitative results quickly, and (iii) is cost-effective using standard hematoxylin and eosin (H&E) stained tissue sections. To address all these needs in a deep learning approach, we utilized the convolutional encoder-decoder network SegNet to train the specialized network IMFSegNet allowing to accurately quantify the spatial distribution of IMF in histological sections. Our fully automated analysis was validated on 17 H&E-stained muscle sections from individual sheep and compared to various state-of-the-art approaches. Not only does IMFSegNet outperform all other approaches, but this neural network also provides fully automated and highly accurate results utilizing the most cost-effective procedures of sample preparation and imaging. Furthermore, we shed light on the opacity of black-box approaches such as neural networks by applying an explainable artificial intelligence technique to clarify that the success of IMFSegNet actually lies in identifying the hard-to-detect IMF structures. Embedded in our open-source visual programming language JIPipe that does not require programming skills, it can be expected that IMFSegNet advances muscle condition assessment in basic research across multiple areas as well as in research fields focusing on translational clinical applications.

1. Introduction

Assessment of muscle condition is a target of interest in various research areas including biology, medicine, psychology, exercise science, biomechanics, engineering, biochemistry as well as food industry. In addition to the widespread analysis of changes in muscle fibers, the analysis of fat infiltration is increasingly becoming a focus of research. The amount of fat in muscle tissue is a crucial factor in assessing the performance and condition of a muscle. Intramuscular fat (IMF) can not only arise in the context of diseases but also because of fibrosis [1] induced by muscle damage. Additionally, muscle fibers can be replaced by fat due to necrotic degeneration [2,3]. The accumulation of IMF above a certain level results in the muscle being unable to maintain sufficient tension [4,5]. In contrast, from the perspective of food industry, a certain amount of IMF is a sign of meat quality [6,7]. Since the various research fields view muscle condition from different perspectives in diverse mammals, there exists a wide range of methods by which muscle condition is examined qualitatively and quantitatively.

IMF analyses have previously been performed based on magnetic resonance imaging (MRI) and computer tomography (CT) [5,8–14], as well as analyses of whole muscles based on decellularization [15], histological examinations of tissue sections with perilipin fluorescence [15] or Oil Red O staining (ORO staining) [11,16–19] and Harris hematoxylin & eosin staining (H&E staining) [16–21]. In some cases, combinations of

* Correspondence to: Leibniz Institute for Natural Product Research and Infection Biology – Hans Knöll Institute, Adolf-Reichwein-Straße 23, 07745 Jena, Germany.
E-mail address: thilo.figge@leibniz-hki.de (M.T. Figge).

https://doi.org/10.1016/j.csbj.2023.07.031
Received 14 April 2023; Received in revised form 21 July 2023; Accepted 22 July 2023
Available online 25 July 2023

macroscopic and histological examinations [11,12] as well as of different histological techniques [17,19] were performed. Accurate quantification of IMF is important for comparable assessment, as emphasized by Ogawa et al. [22] in the context of analyzing IMF by MRI. Since MRI and CT are relatively expensive approaches [15] and as taking micro-biopsies is becoming a common practice [23], analysis of IMF from histological sections is a comparably cost-effective approach for assessment of muscle condition. Moreover, histological analyses can be used to further verify MRI and CT results, when using an ordinary transmitted light microscope with a camera.

The application of ORO or H&E staining is used to perform histological analysis of IMF at a qualitative level, as well as at a quantitative and semi-quantitative level. Applying ORO staining yields IMF in a highly saturated red tone (Fig. 1A blue arrow) with the surrounding muscle tissue in a purple-blue to pale pink tone (see Fig. 1A). In contrast, the easy-to-apply and cost-effective standard H&E staining yields pink-red coloring of muscle fibers and a sallow coloring of IMF cells with a characteristic orbital shape (see Fig. 1B green arrow). Cracks and holes are not colored in both stains (see Figs. 1A and 1B orange arrow). Experts can distinguish pale IMF areas, from fatless holes and cracks or structures that could be confused with fat cell, such as macrophages, because of its typical orbital shape, in an H&E-stained muscle section. The determination can be facilitated and supported by the analysis of adjacent muscle sections stained with ORO.

In particular, analyses based on a scoring system defined by experts [11,12,19], can be highly subjective even when used by independent scientists, as has been reported in Kim et al. [19] and is generally known as inter-observer variability [24,25]. These considerations are a strong motivation to perform the image analysis in an automated manner, for which various candidate approaches should be compared.

Automated IMF analysis has previously been performed using the image processing tool ImageJ [26], where images of tissue sections were segmented based on color or gray-level as a standard feature to distinguish different regions [16,18,20,21]. The challenge here is to avoid false-positive assignments of cracks and holes or structures that could be mistaken for fat cells. Although regions with IMF appear very clearly in the case of ORO staining, its application is accompanied with severe disadvantages. Application of ORO staining is more time-consuming, twice as expensive as H&E staining, and samples can be challenging to handle. The latter issue is due to the fact that, albeit the analysis by color becomes simpler as such, automated analysis will be error-prone due to observable staining-related artefacts. Speckles or stain bleeding (see Fig. 1A red arrow) can lead to areas which may erroneously be characterized as fat, as was pointed out by Biltz und Meyer [15].

Overall, we aimed to find an automated image analysis approach that meets the following objectives: (i) minimization of subjectivity, (ii) rapid provision of accurate quantitative and qualitative results, and (iii) cost efficiency using standard H&E-stained tissue sections. In order to achieve this, we performed a comparative evaluation of selected image analysis workflows. From all the image analysis workflows that we have tested, the challenging-to-detect features of IMF regions were most accurately identified by a deep learning-based approach. We refer to this neural network as IMFSegNet, because it was trained on the well-known SegNet architecture [27]. However, deep learning approaches are known to be data-intensive and, in order to learn several millions of parameters, they require manually annotated data sets that are time-consuming to prepare. Therefore, we also attempted to achieve our objectives using a classical image analysis approach based on label-free k-means clustering [28] of color-based intensity features, which only requires the cluster number to be specified. Next, we used the interactive machine learning-based tool ilastik [29] to incorporate additional shape- and texture-based features from the muscle sections. The number of parameters for intensity-, edge-, and texture-based features is of the order of 100 and the parameters are adjusted during training on just a few manual annotations that can be specified within a few minutes. In contrast, deep learning-based workflows demand massive manual annotations to fit a convolutional neural network (CNN) with more than one million parameters. These manual annotations are required to be of high quality and this usually amounts to a work load of several hours to days. The advantage of these workflows is that they are capable of detecting fine-grained features in a data-driven manner during their training process. We tested this by applying a transfer learning strategy, where the neural network Cellpose [30], which was previously trained on different types of cells, was fine-tuned to recognize fat cells. However, our comparative assessment revealed that, in order to not only achieve the best quantitative results for fat cell detection but to also ensure a generalizable approach that meets our objectives, we had to train an artificial intelligence from scratch. This was successfully realized by training IMFSegNet on histological sections of muscle tissue using the SegNet architecture [27]. While the exact process of decision making in deep neural networks cannot be reconstructed and analyzed in detail due to millions of model parameters, several approaches have been developed in recent years to create explainable artificial intelligence [31,32]. In order to shed light on the opacity of IMFSegNet regarding the identification of IMF in tissue sections, we applied the Gradient-weighted Class Activation Mapping (Grad-CAM) [33], a technique to visually elucidate the decision-making process of a CNN. This workflow required precise and time-consuming manual annotation and based on this learned in the order of 10 million parameters during training with the advantage of producing generalized and robust results on unseen data. All workflows were implemented in our recently introduced Java Image Processing Pipeline (JIPipe: http://www.jipipe.org). JIPipe is an open-source visual programming language for easy-access pipeline development [34] that allows extensive comparisons of a wide variety of image analysis methods we used to segment IMF regions in H&E-stained muscle tissue.

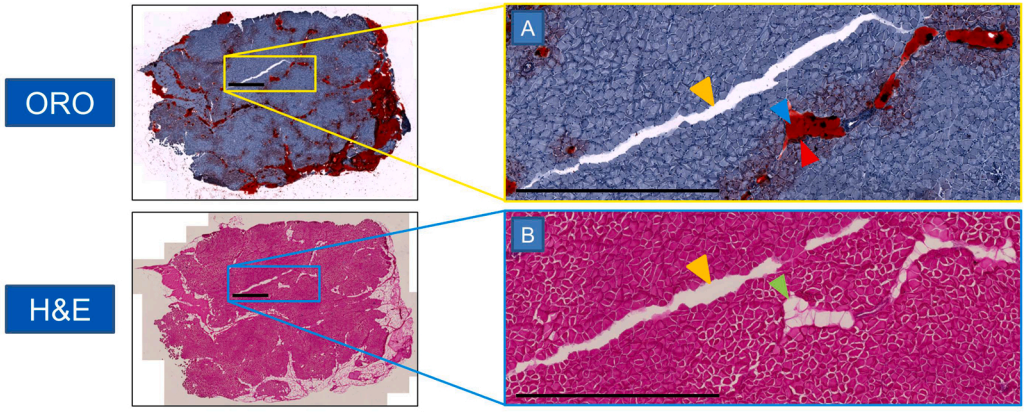

Fig. 1. Example of Oil Red O (ORO) and hematoxylin and eosin (H&E) stained sheep muscle section. ORO (A) was performed on tissue sections from the adjacent H&E stained (B) tissue sections. The orange arrows (A and B) indicate the cracks which are not colored in both stains. The blue arrow (A) denotes the highly saturated red tone of the ORO stained fat cell. The green arrow (B) denotes the intramuscular fat cells with the characteristic orbital shape in the H&E stained tissue section. The red arrow (A) indicates staining-related bleeding, which can lead to increased false-positive evaluation of IMF. Scalebars correspond to 1 mm.

2. Methods

2.1. Data collection and staining

To avoid the use of human or animal muscle biopsies from living individuals we have used fresh muscle samples from animal cadavers for the development of our muscle tool. The posterior cricoarytenoid muscle and the vastus lateralis muscle were excised from 17 fresh cadavers of adult female Merino sheep (aged between 4 and 5 years, weighing 70–110 kg). The muscles were placed on a small cork plate with tissue embedding medium (Leica®, Germany), frozen for 20 seconds in melting isopentane precooled in liquid nitrogen and stored in liquid nitrogen at −80 °C. Muscles were then cut into 10 μm slices using a freezing microtome (Leica CM 3050 S, Leica®), placed on glass slides (HistoBond® adhesive slides, Paul Marienfeld GmbH & Co. KG) and stored at −21 °C. We used 17 muscle sections, thawed for 30 minutes, for standard H&E-staining [35]. For the ORO staining, 17 adjacent slices were thawed. The sections were rinsed with distilled water followed by 60 % propanol afterwards. Then, the sections were stained with ORO solution (Oil red O staining solution, Sigma-Aldrich®, ORO solution) for 15 minutes, rinsed in 60 % propanol shortly and in distilled water again. Finally, the slices were covered with aqueous glycerol gelatin (Kaisers Glycerin-Gelatin, phenol-free, Carl Roth ®). It is crucial that the coverslips are carefully placed on the sections and not pressed on, which makes this staining approach very sensitive and error-prone for studies with large sample sizes.

2.2. Imaging and annotation

The stained sections were imaged using a Zeiss AxioScan 7 with a magnification of 20x and a resolution of 0.442 μm x 0.442 μm per pixel. We used this resolution because a higher resolution would require downsampling the fat cells due to their size, and because it is not possible to use even larger patches on a common GPU. In the following, the IMF areas of 17 H&E images were manually annotated in ImageJ [26] using the "Wall Tool" and the "Selection Brush Tool" with the help of adjacent samples stained with ORO serving as additional references for a precise annotation (see Figs. 1A and 1B and Suppl. Fig. 1). The manual annotations of these two experimentalists on H&E sections were then quantitatively compared with the ORO stained samples. This comparison was made based on the computed Pearson correlation coefficient and the fat fraction that correlated better with the ORO stained sections. This is detailed shown in Suppl. Fig. 2.

2.3. Performance measures

The pixel-overlap between images that have been automatically analyzed and their manually annotated counterparts was evaluated by the following performance measures: (i) Dice coefficient (DC), (ii) precision (P), and (iii) recall (R), where the latter two account for the error

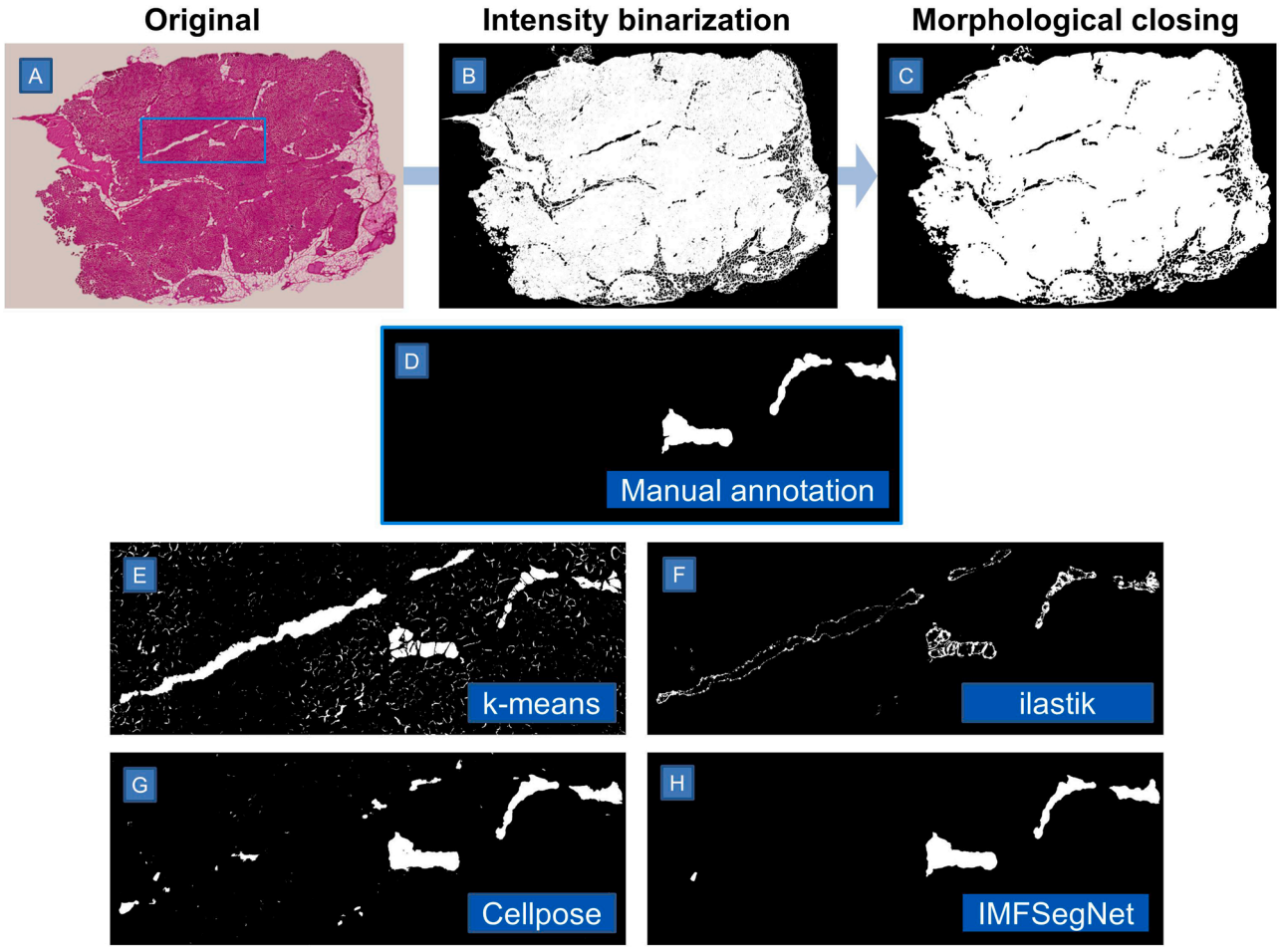

Fig. 2. Segmentation of whole muscle tissue. After intensity-based thresholding (B) of the stained tissue section (A), small perforated structures of the muscle tissue were closed with a subsequent morphological closing operator (C). The results were finally used along with the manual annotations (D) and the fat segmentations (E-H) as the final segmentation of the muscle tissue, from which the fat fraction was determined. While k-means clustering (E) generally tends to over-detect fat and, on the other hand, miss the fibers between fat cells, ilastik (F) detects fat predominantly at the boundary of fat tissue and ruptures, but tends to not capture the fat cell in its entirety. While Cellpose (G) yielded more false-positive errors, the IMFSegNet (H) overlaps the most with the manual annotations (D).

due to false positives (FP) and false negatives (FN), respectively. In order to provide a single score for the IMF fraction between the manual annotations and the model predictions we determined the (iv) Pearson's product-moment correlation coefficient [36]. In addition to measuring the pixel-wise overlap between the manual annotations and the model predictions, we also calculated the (v) Hausdorff distance [37] (see Suppl. Fig. 3), which determines the maximum distance between a manually annotated pixel as being IMF and the nearest predicted pixel to be IMF within an image. All measures related to pixel overlap utilized here consider (non-) overlap between two sets, where the two sets consist of the manual annotations and the model predictions. The complete overlap between both sets is defined as true positive (TP), while the over-detection of the predictions, which are not present in the manual annotations, is defined as FP. In the opposite case, all non-detected values of the manual annotations are defined as FN. The Dice coefficient, precision and recall are defined as

$$DC = \frac{2TP}{2TP + FP + FN}, P = \frac{TP}{TP + FP}, R = \frac{TP}{TP + FN}$$

The (iv) Pearson's product-moment correlation coefficient also provides a test for paired samples which gives a p-value when the correlation is significantly different from zero across all samples, meaning that the two variables are significantly correlated with each other. In the context of IMF analysis, this implies that the calculation of correlation involves using as input both the IMF fraction predicted by the model and the manual annotations. These two factors exhibit a significant correlation within a muscle section. The significance levels were set to: *= $p < 0.05$, ** = $p < 0.01$ and ***= $p < 0.001$. Since the (v) Hausdorff distance measure is not symmetrical, we have used the maximal value of both directions of the respective sets as follows:

$$d_H(X, Y) = \max\left\{ \sup_{x \in X} d(x, Y) \middle| \sup_{y \in Y} d(X, y) \right\}.$$

In this case, a subwindow of a whole tissue section containing manually annotated IMF, which we refer to here as a patch, has the size of 512×512 pixels, where the maximum possible distance is $\sqrt{512^2 + 512^2} = 724.1$ pixels, and by which all determined values were normalized to obtain values between zero and one. The patch size of 512×512 pixels was chosen to accommodate the tradeoff between providing the greatest possible context and information flow for the CNN and the computational and storage capacities.

2.4. Segmentation of whole muscle tissue

In order to determine the IMF fraction for each muscle individually the muscle tissue had to be segmented as a whole. This was realized using the visual programming language JIPipe [34], which is based on the image processing tool ImageJ [26]. To obtain a binary image that distinguishes between tissue and background, we converted the original image (see Fig. 2A) into a color space typical for hue segmentation. Subsequently, the tissue could be segmented sufficiently well by visual inspection using a threshold value of 15 % in the saturation channel (see Fig. 2B). It can be concluded that the light background generally has no saturation within its corresponding hue. To remove remaining artifacts, we used Remove Outliers 2D from ImageJ, which is designed to correct for dead pixels. This process replaces a pixel with the median of the surrounding pixels if the intensity value deviates from the median by more than a certain threshold. The number of surrounding pixels within a so-called kernel, which was used to calculate the median, is here 31 pixels in radius. This threshold value ensures the operator to just eliminate noise, because it corresponds to the size of objects that are less than 12% in size compared to average fat cells. To close holes in perforated muscle tissue that are smaller or equal than 5% of the average fat cell size, we applied morphological closing (see Fig. 2C) with a disk shape and a kernel size radius of 13 pixels. Finally, the total area of each muscle tissue was quantified. For subsequent calculation of the IMF fraction in muscle tissue and to minimize FP induced by the different image analysis workflows in this study, we determined the convex hull of the final tissue mask and preserved only the predictions that fell within this convex hull.

2.5. Automated fat segmentation in muscle tissue

To perform the fat segmentation using a white-box approach, where one can accurately track how the final fat-non-fat discrimination comes about, we first segmented the stained sections according to their color, using a hue- and saturation-based image analysis approach. For this purpose, we converted each image to CIELAB color space and applied unsupervised k-means clustering [28] via the python library "scikit-learn" [38] with three clusters to all tissue sections on top of it (see Fig. 2E). Since the results were not satisfactory, we used the interactive machine learning tool ilastik [29], where we used the manual annotation of one entire tissue section for model training and then applied it to all other sections (see Fig. 2F). The limitation of using one sample for

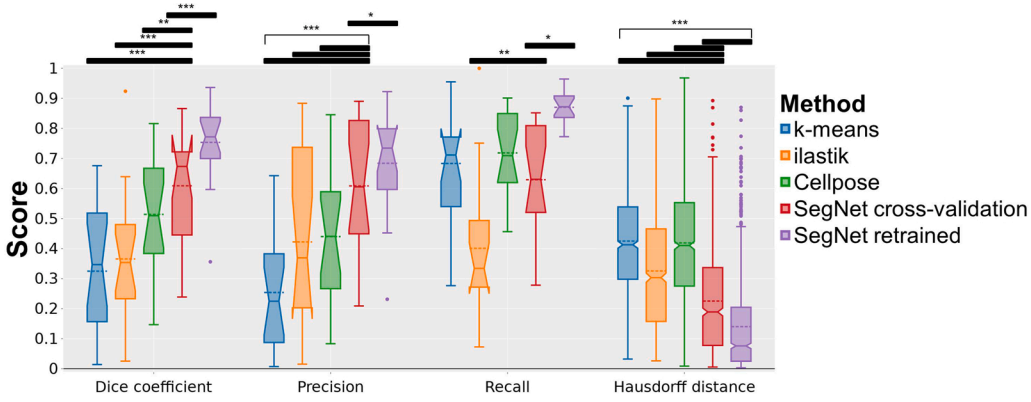

Fig. 3. Method comparison by validation in terms of the pixel-based performance measures Dice coefficient, precision, and recall. Moreover, we computed the Hausdorff distance, which indicates the maximum distance of a manually annotated pixel to a pixel classified as fat within a subwindow with a size of 512×512 pixels, which we refer to as a patch. These patches were used to train the deep learning-based techniques Cellpose and SegNet. Each value of the Hausdorff distance was normalized by the maximum possible distance of 724.1 pixels. In contrast to the pixel-based performance measures, the better the performance the closer the Hausdorff distance is to zero. We compared all methods across the corresponding metrics using an unpaired t-test, indicating only significant differences between SegNet in the context of cross-validation and all other respective methods. This was done to verify that the SegNet is most capable of generalizing on unseen test data. The SegNet achieved significantly better results than the approaches k-means and ilastik across all measurements. Compared to the deep learning-based approach Cellpose, SegNet achieved significantly better scores in terms of precision and Hausdorff distance. For all metrics, SegNet retrained on all available samples produced the significantly best results compared to all other methods. Stars correspond to: *=p-value < 0.05, ** =p-value < 0.01, ***=p-value < 0.001.

training was chosen due to the computationally intensive training duration. The fat segmentation in ilastik was performed with the Pixel Classification method using all intensity, edge and texture features after a previous smoothing of the image using a Gaussian filter and the respective sigma-values of 0.7, 3.5, and 10 pixels. Next, we applied deep learning-based approaches capable of learning fine-grained features in a data-driven manner. Since the image size for each of the seventeen histologically stained muscle sections is on the order of 10^4 x 10^4 pixels and the application of deep learning at such an image size is limited by hardware components, we focused on the manually labeled fat regions. These manually annotated fat regions account for an average of 4.3% ± 4.2% of the total tissue. For this purpose, we extracted all the fat regions into binary patches of 512 × 512 pixels each, with the foreground referring to the fat and the background referring to the rest. This patch size was chosen to optimally utilize the computing and hardware capacities. Since Cellpose [30] has recently become a generalizable and established deep learning-based approach with outstanding segmentation results of different cell types, we retrained it on all 1629 patches from within JIPipe, specifying an average diameter of 210 pixels, which corresponds to 93 µm, i.e. the average measured diameter of the manually annotated regions (see Fig. 2G). We trained the model to a maximum of 500 epochs with a batch size of 32 images. To further reduce the imbalance of false-positives and false-negatives, we used SegNet [27], a deep convolutional encoding-decoding architecture for semantic segmentation (see Fig. 2H). To test whether CNN can be applied robustly and generalized to unseen data, we first performed an eight-fold cross-validation stratified on individual sheep (see Suppl. Fig. 4). We trained each model with a maximum of 1000 epochs, including a termination criterion, in case the model did not perform better for 300 epochs on a randomly selected validation set of 20 % within the training dataset and a batch size of 32. To have more data available during training, we used standard data augmentation techniques [39] such as rotating, flipping, and zooming, which increased the size of the training dataset by a factor of five. To facilitate the model's learning to discriminate between the underrepresented fat tissue and the remaining tissue, we assigned balanced weights to the two labels during training. The output of the model is a probability value per pixel, where the value of one represents the maximum probability of being identified as fat, or zero represents the minimum probability of being classified as fat tissue. Finally, to obtain a binary image, the Isodata [40] algorithm was applied to the model prediction of each image. In order to obtain a model prediction on the entire tissue section, a sliding window process is applied over the whole image. It is known that the Unet [41] architecture is prone to errors during decoding, also called upsampling, and can produce artifacts at the outer edges of each patch. To overcome this issue, we used the overlap-tiles function, which considers only the center per tile for the final image prediction. After cross-validation of SegNet, the same approach was retrained again on all samples, as it was done for Cellpose. This retrained SegNet model was then used as IMFSegNet workflow for the more comprehensive fat analysis of all laryngeal

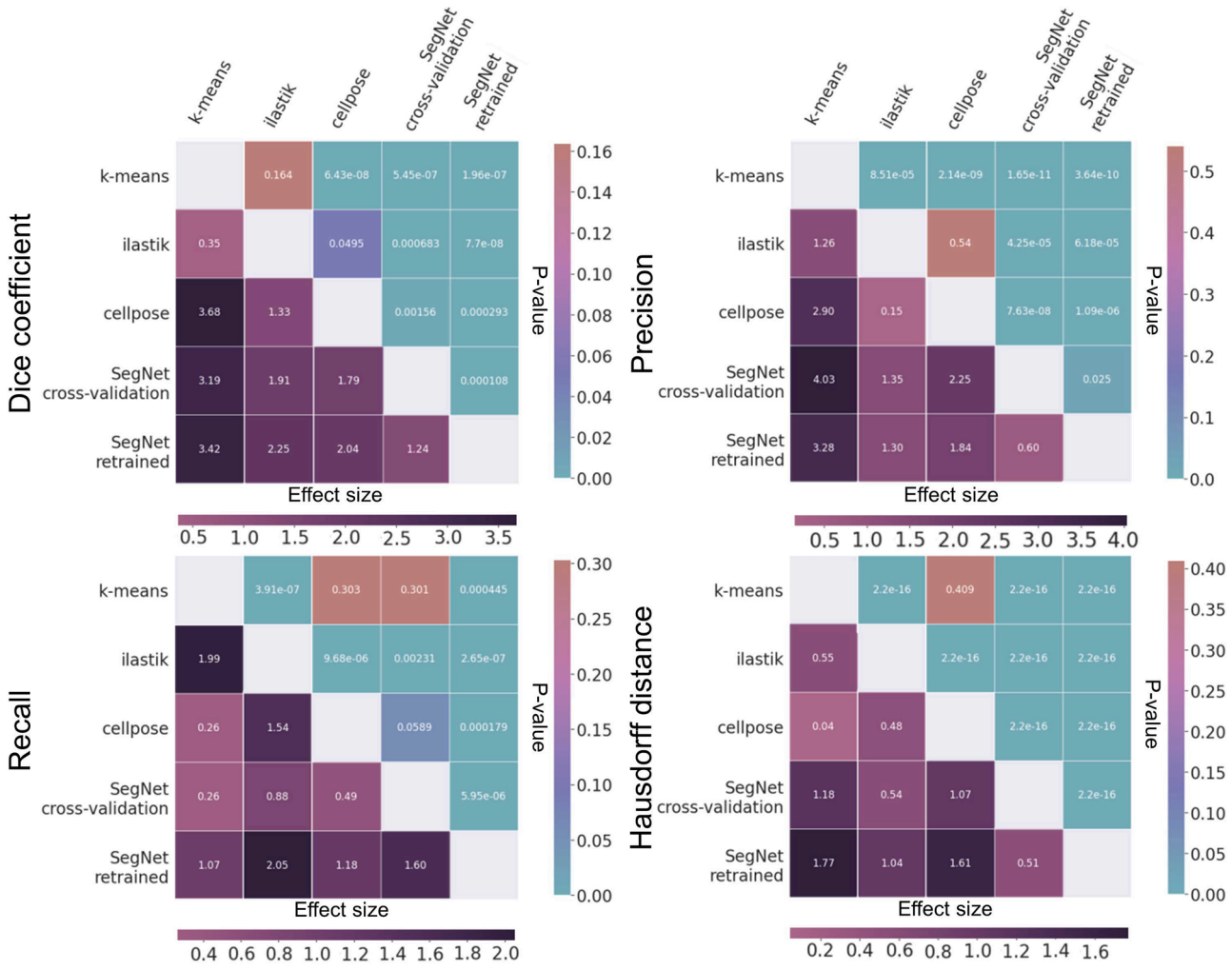

Fig. 4. Statistical evaluation between all applied methods across all metrics. The matrices give the results of the statistical tests for the metrics of the dice coefficient, precision, recall and the Hausdorff distance. The upper triangular matrices refer to the p-values and the lower triangular matrices to the corresponding Cohen's d effect sizes.

muscles in the 17 sheep included in this study. To compare not only a pixel-based evaluation but also the relative detected IMF fraction per tissue sample for all methods we calculated Pearson's product-moment correlation coefficient between the manually annotated tissue sections and the automated segmentations from the analysis per image.

2.6. Incorporation of explainable artificial intelligence

To gain more insight into the IMFSegNet decision-making process, we used Grad-CAM [33], a technique to make the decision-making process more transparent with visual representations. Grad-CAM uses the gradients of the existing labels expressed in the last convolutional layer to create an approximate localization map, which highlights the image regions that are important for predicting a particular label, i.e. regions of IMF. We applied Grad-CAM within a jupyter notebook in python to the tissue section seen in Fig. 2 (see Suppl. Fig. 5). The implementation was done in such a way that it is easy to extend the analysis to further samples (see jupyter notebook "SheepFat_2_SegNetSegmentaion" in the github repository).

2.7. Data and code availability

All material, consisting of the entire JIPipe project, Python scripts including jupyter notebooks and ilastik files are available for download at:

https://github.com/applied-systems-biology/sheepfat.

All data is available for download here:

https://asbdata.hki-jena.de/PraetoriusEtAl2023_ComputStructBiotechnolJ.

The versions of the used tools such as JIPipe is 1.74, for the k-means clustering in sklearn 1.0.2, for ilastik the version 1.3.3, Cellpose 0.7.3 and the IMFSegNet was developed using tensoflow-gpu version 2.5.1.

2.8. Statistical calculation

To determine significant differences between the methods used for the (i) - (iii) pixel-wise measurements and the (v) Hausdorff distance, they were compared using a paired t-test.

All p-values were defined as being not significant for $P > 0.05$ and significant for $P \leq 0.05$. When values were defined as significant, effect size was determined by calculating Cohen's d [42] using the R library "effectsize". The ranges of effect size magnitudes are referred to as *negligible* for $|d| < 0.2$, small for $|d| < 0.5$, medium for $|d| < 0.8$ and large for $|d| \geq 0.8$.

2.9. Ethics statement

The study was performed in accordance with the European and German animal welfare regulations. The muscle samples of 17 fresh cadavers of adult female Merino sheep were analyzed, euthanized in the context of other projects.

3. Results

All image analysis workflows that were compared in this study have been validated on the 17 H&E-stained sheep muscle slices to select the workflow that not only provided the best IMF segmentation in terms of pixel-wise overlap with the ground truth, but also yielded the best IMF fraction per muscle tissue relative to the annotations. To evaluate the generalization of IMFSegNet to unseen test data within cross-validation runs, all workflows were compared using various metrics. Additionally, the statistical significance of IMFSegNet in cross-validation was compared to that of all other workflows. All results related to the pixel-wise metrics – i.e. dice coefficient, precision and recall – as well as the Hausdorff distance are summarized in Fig. 3. The corresponding statistical tests are given in Fig. 4 and the results for Pearson's product-moment correlation coefficient are given in Table 1.

The classical image analysis workflow relying on a color-based k-means clustering performed well in terms of a median recall of 71 %. In contrast, however, the median precision is only 22 %, indicating that the method does not miss IMF, but over-detects IMF to a high degree. The two measurements combined resulted in a median Dice coefficient (DC) of 35 %. The high degree of IMF over-detection was also reflected in Pearson's correlation coefficient of 0.39, which is not significantly different from zero ($p = 0.12$), implying that we cannot speak of a correlation here. The interactive machine learning-based approach ilastik [29] exhibited a more balanced performance in terms of median precision (37 %) and recall (33 %), but overall resulted in a DC of 35 %, which is as low as for the classical image analysis workflow. In terms of detecting the correct IMF fraction, this method performed even worse with a correlation coefficient of 0.18 ($p = 0.5$) implying that the ilastik approach did not get beyond a result better than flipping a coin per pixel. Given these quantification results (see Fig. 3), in combination with the visually representative IMF segmentation (see Fig. 2), we saw the need for workflows that are capable of examining higher-order structure and texture features of muscle tissue in order to achieve our objectives.

In the next step, we applied deep learning-based workflows that learn image features in a data-driven manner. Using transfer learning with Cellpose yielded a DC of 51 % based on a recall of 71 % and a precision of 44 %. These quantification results implied that, like for the k-means clustering workflow, IMF was over-detected rather than being missed in the segmentation. In contrast, the IMF fraction prediction of Cellpose correlated significantly (p-value $\ll 10^{-6}$) with a Pearson's correlation coefficient of 0.91, implying that this workflow generally detected the relative amount of IMF within a tissue section quite well. To further improve the results, the SegNet architecture was used and an eight-fold cross-validation (see Suppl. Fig. 4) was performed, which allows to evaluate the generalization and robustness of the model that we refer to as IMFSegNet. The evaluation showed a median DC of 67 %, a precision of 61 %, and a recall of 63 %, which is indicative for having reached a balance between false-positives and false-negatives. Moreover, in terms of DC being the harmonic mean between precision and recall, IMFSegNet achieved significantly better results than all other tested workflows, i.e. k-means clustering (p-value $\ll 10^{-6}$, Cohen's d [42](d) = 3.19), ilastik (p-value $\ll 10^{-3}$, d = 1.91) and Cellpose (p-value = 0.0016, d = 1.79). The IMF fraction detected by the IMFSegNet model on all test samples correlated significantly (p-value $\ll 10^{-9}$) with a Pearson's correlation coefficient of 0.97. Using the overlap tiles feature described in the Methods section, we were even able to increase the DC median by 3 %. After evaluating the IMFSegNet model with cross-validation on unseen test data, it remained to be noted that the results of this approach were better than with all other compared workflows. Based on this best practice, we retrained it on all 17 manually annotated tissue sections, because we anticipated that the best generalized predictions can be expected from a model trained on all available tissue sections and whose parameters have been optimized. Retraining the IMFSegNet led to even better pixel-based quantitative results, specifically a median DC of 77 %, precision of 73 %, and recall of 82 %. Moreover, the retrained IMFSegNet exhibited the highest significance (p-value $\ll 10^{-14}$) in the Pearson's correlation coefficient (0.99) compared to all other tested image analysis workflows (see Figs. 3 and

Table 1
The Pearson's moment correlation coefficient with corresponding p-value for each of the methods used. Bold value highlights the best result.

Method	Pearson correlation	P-value
k-means clustering	0.39	0.12
ilastik	0.18	0.50
Cellpose	0.91	$<<10^{-6}$
SegNet cross-validation	0.97	$<<10^{-9}$
SegNet retrained	0.99	$<<10^{-14}$

4).

Computation of the Hausdorff distance d_H (see Suppl. Fig. 3) revealed that both the k-means clustering workflow, with a d_H median of 0.41, and the deep learning based Cellpose, with a d_H median of 0.41, performed similar and turned out to yield the highest d_H values. A random forest pixel-classifier, which is implemented in the interactive program ilastik, ranked this workflow significantly better with a median Hausdorff distance of 0.3 (p-value $\ll 10^{-15}$, d = 0.55 in comparison with k-means clustering and d = 0.48 compared to Cellpose). The IMFSegNet in the scope of a cross-validation achieved a d_H median of 0.19 and by that significantly better results than k-means clustering (p-value $\ll 10^{-15}$, d = 1.18), ilastik (p-value $\ll 10^{-15}$, d = 0.54) and Cellpose (p-value $\ll 10^{-15}$, d = 1.07). As for the pixel-based metrics, the retrained IMFSegNet performed best with regard to the median Hausdorff distance. Specifically, even if the pixel-wise overlap in DC, precision, and recall is not perfect, nevertheless the prediction is generally close to ground truth with a median value of the Hausdorff distance of only 0.08 corresponding to 55 pixels or 27.5 μm.

The application of Grad-CAM reveals by visual inspection that the IMFSegNet has the highest gradient activation where corresponding fat was manually annotated. In this regard, there is an indicator that the IMFSegNet decision process for IMF is based on the characteristic orbital shape with the texture-typical transparent interior of the fat cell and fine contours of the membranes between the individual cells (see Fig. 1B green arrow and Suppl. Fig. 5).

4. Discussion

4.1. Comparative assessment of various image analysis approaches

Automated and objective analysis of IMF under cost-effective conditions for H&E staining proved to be a challenge. Since fat appears in these sections very similar to the background color, the white-box approach using an unsupervised k-means clustering [28] workflow on color intensities did not yield high performance measures. Although ilastik [29] has generally proven to be a very powerful machine learning-based approach that requires only a few annotations for image segmentation and analysis [43,44], it did as well fail to provide highly accurate results for IMF segmentation.

To incorporate high-level features that are extracted in a data-driven manner, we needed to rely on deep learning workflows. In this context, Cellpose [30] has proven to be a generalist approach for segmenting a wide variety of cell types and has achieved outstanding results [45]. Training a model on the data from this study, Cellpose was found to improve segmentation in general, but still revealed weaknesses with regard to over-detection of IMF in the tissue. For this reason, going beyond Cellpose, we searched for a sophisticated workflow that would accommodate the demanding requirements for IMF segmentation in tissue sections and chose to build on the SegNet [27] architecture. SegNet was specifically designed for semantic segmentation and is capable of learning high-level features. We trained this neural network to identify IMF in tissue sections and referred to the resulting model as IMFSegNet, which yielded accurate quantification of IMF in unseen data of muscle sections. Although the Dice coefficient appeared low with a median of 67 %, it should be noted that the IMFSegNet was the only model that was able to achieve balanced results in terms of false-positives by the assessment of precision (61 %) and false-negatives by the assessment of recall (63 %). In addition, the correlation of IMF fraction between the manual annotation and the model prediction was highest. The evaluation with the Hausdorff distance lead to the conclusion that, despite the fact that pixel-wise overlaps of the IMFSegNet do not exceed a desirable value of 90 %, as is usually demanded from deep learning methods, the detected IMF was always found in the immediate vicinity of IMF in the ground truth. Morover, compared to Cellpose, ilastik and k-means clustering, the use of IMFSegNet in the context of cross-validation yielded the most significant quantitative results (see Fig. 3 and Fig. 4). This observation prompted us to proceed with the IMFSegNet retrained on all samples, which overall produced the best quantitative and visual results on the 17 manually annotated tissue sections (see Fig. 2), making it the most suitable of all investigated workflows for the challenging analysis of tissue sections. Although the use of Grad-CAM [33] is only indicative of the decision-making process by CNN, the use of this technique for explainable artificial intelligence has contributed to gaining better insight on the focus of black boxes such as IMFSegNet regarding the decision-making process of identifying hard-to-detect IMF regions as such.

The release as well as combine advanced prompt-based chatbots such as ChatGPT with models that generate images from text descriptions, such as Dall-E, give their users an idea of how powerful artificial intelligence-based methods have become. An appropriate adaptation of such diffusion models for automated and high-precision biomedical image analysis could once again raise the CNN-based achievements to a new level and further improve the results. Meta provides first steps in this direction with its Segment Anything Model (SAM) tool [46].

Although the IMFSegNet has demonstrated its generalizability in cross-validation, it may not be able to work when applied to other H&E stained sections if there are too many color differences due to, for example, a modified staining protocol. In addition, the model may have potential issues detecting fat in H&E stained muscle sections from other individuals where the fat cells differ substantially in size, shape and texture from those on which the model was trained in this study In order to adjust the model so that it can perform an analysis on the new data, it is necessary to make new manual annotations and to retrain the model itself. However, the latter requires specific hardware or outsourcing the data to cloud-based systems where data security may be questionable. Retraining, often known as transfer-learning, can therefore prove difficult and another bottleneck to further generalizability of the model. At this stage, it cannot be excluded that further improvements in the segmentation accuracy at the pixel-level may be achieved by using the new technology of VisionTransformers [47], or by extensive and advanced preprocessing. While it will be left to future investigations to optimize the preprocessing step, it should be pointed out here that this can be realized within JIPipe [34], which provides a synergistic opportunity by the combination of classical image processing and deep learning methods.

4.2. Cost-effective analysis and objective quantification of fat infiltration in muscle tissue

It should be emphasized that IMFSegNet does not confuse cracks or other structures that could be identified as fat cells in tissue sections with the occurrence of IMF (see Fig. 2), in contrast to other H&E methods mentioned before [20,21]. In addition, the 17 cross-sections included in this study are characterized by a large variation of different staining intensity, which can be seen in Suppl. Fig. 4. Thus, we have attempted to provide the best possible data basis for high quality and generalizing model predictions.

Since the fiber type composition is similar in all mammals [48–50], IMFSegNet developed on sheep muscle can be used for accurate IMF analysis on all mammals, including humans.

But IMFSegNet is more than just an accurate IMF analysis. IMFSegNet complements expensive imaging modalities such as MRI or CT, for which deep-learning approaches have also been used to quantify fat infiltration and occupancy [51–54], but have lower resolution than histological imaging modalities. IMFSegNet also circumvents fluorescence sample preparation and imaging, which is more than four times as expensive and more time-consuming by a factor 45, as well as ORO sample preparation and imaging, which is twice as expensive and twice as time-consuming. Microscopy using fluorescence labeling does often not only have more costly aspects, it can also have causal effects on the experimental outcome [55]. In addition, a transmitted light microscope

with a camera would also be easier to use and less expensive than more complex fluorescence microscopes and can be found in many laboratories, even in a clinic.

Our automated deep learning-based workflow can measure H&E stained-muscle biopsies of various sizes, and in combination with micro biopsy [23], can be applied in the context of minimally invasive sample acquisition. Furthermore, the ready-trained IMFSegNet can be run on standard computational hardware allowing to perform muscle analysis on the spot and does not need to be shipped to a pathology laboratory [56]. Running a prediction of IMFSegNet requires on the CPU of an average consumer workstation (Intel Core i7–8700 K, 32 GB RAM) a computation time of about 320 seconds per muscle slice (10k * 10k pixel size). Additionally, snap-frozen muscle biopsies can be used, as is common practice for muscle disease diagnosis [57]. For example, it is not only recommended to constantly analyze damaged [58,59], aged [60] and exercised muscles, but also patients with muscular diseases [56,57,61,62], obesity or diabetes [63], in order to monitor treatment effectiveness and overall health condition. For instance, it has been pointed out before that in resource-limited settings such as developing countries, muscle biopsy may be the preferred method for diagnosing muscular dystrophy. This is mainly due to the reduced access to MRI or high costs of genetic testing for unique mutations impairing genetic diagnosis [62]. Moreover, ethical conflicts, as may result from genetic testing for the diagnosis of muscle dystrophy, can be avoided because the AI analysis from H&E muscle cross-sections alone cannot re-identify the patient. At the same time, physicians may still retrieve important and useful information for an individual therapeutic application. Furthermore, this technique could as well be valuable to the food industry and farmers who want to determine the quality of meat by analyzing the IMF, as evidenced by studies exploring the "official beef marbling standard" or assessing chicken meat quality using histological methods [64,65]. In all these situations it is desirable to have a cost-effective and objective quantification as provided by IMFSegNet.

5. Conclusion

The application of deep learning is applicable in many biomedical fields, such as artificial histochemical staining [66], fast and accurate cancer detection [67], and tremendous time efficiencies via automated animal behavior tracking [68]. These advantages and achievements have made deep learning the best choice to analyze IMF in H&E-stained muscle sections in an accurate, objective, performant and automated manner. In conclusion, the applied artificial intelligence-based IMFSegNet combined with the visual programming language JIPipe [34] enables objective, automated, and accurate quantification of IMF in muscle sections within seconds that can be performed without requiring programming skills. We envisage IMFSegNet to be a step towards making conventional ordinal-scaled and subjective characterization by a pathologist obsolete in the future. In fact, while artificial intelligence may take over the routine work, the future role of pathologists may be more supervising the automated process. To our knowledge, standardized tools for IMF analysis in H&E-stained muscle sections do not exist today, which supports the use of IMFSegNet for various complex muscle fat analyses on cost-effective sample preparation.

Author contribution

JPP, CMS and MTF designed this study. KW and DA did the manual annotations. JPP performed the image analysis. JPP and KW performed the statistical analysis. JPP, KW and MTF performed the interpretation of the calculations. JPP, KW and MTF wrote the manuscript and CMS, DA and MTF supervised the work.

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.

Acknowledgement

The authors would like to thank MED-EL Elektromedizinische Geraete Gesellschaft m.b.H, Austria, the Department of Otorhinolaryngology, SHR Wald-Klinikum Gera, Germany, Division of Phoniatrics, Medical University of Graz, Austria and Institute of Orthopaedic Research and Biomechanics, Centre for Trauma Research Ulm, Germany, for their scientific cooperation and for providing the samples underlying our research. This work was funded by the German Research Foundation (DFG) through the CRC/TR 124 FungiNet with project number 210879364 (Project B4 to MTF). Financial support was also provided by the German Federal Ministry of Education and Research within the funding program Photonics Research Germany, Project Leibniz Center for Photonics in Infection Research, Subproject LPI-BT3, contract number 13N15709.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.csbj.2023.07.031.

References

[1] Sciorati C, Clementi E, Manfredi AA, Rovere-Querini P. Fat deposition and accumulation in the damaged and inflamed skeletal muscle: cellular and molecular players. Cell Mol Life Sci 2015;72:2135–56. https://doi.org/10.1007/s00018-015-1857-7.
[2] Ashley Z, Sutherland H, Lanmüller H, Russold MF, Unger E, Bijak M, et al. Atrophy, but not necrosis, in rabbit skeletal muscle denervated for periods up to one year. Am J Physiol-Cell Physiol 2007;292:C440–51. https://doi.org/10.1152/ajpcell.00085.2006.
[3] Ashley Z, Sutherland H, Russold MF, Lanmüller H, Mayr W, Jarvis JC, et al. Therapeutic stimulation of denervated muscles: the influence of pattern. Muscle Nerve 2008;38:875–86. https://doi.org/10.1002/mus.21020.
[4] Addison O, Marcus RL, LaStayo PC, Ryan AS. Intermuscular fat: a review of the consequences and causes. Int J Endocrinol 2014;Vol 2014:1–11. https://doi.org/10.1155/2014/309570.
[5] Gerber C, Schneeberger AG, Hoppeler H, Meyer DC. Correlation of atrophy and fatty infiltration on strength and integrity of rotator cuff repairs: a study in thirteen patients. J Shoulder Elb Surg 2007;16:691–6. https://doi.org/10.1016/j.jse.2007.02.122.
[6] Miller RK. Chemical and physical characteristics of meat | Palatability. Encycl. Meat Sci. Elsevier,; 2014. p. 252–61. https://doi.org/10.1016/B978-0-12-384731-7.00085-4.
[7] Lonergan SM, Topel DG, Marple DN. Fat and fat cells in domestic animals. Sci. Anim. Growth Meat Technol. Elsevier,; 2019. p. 51–69. https://doi.org/10.1016/B978-0-12-815277-5.00005-6.
[8] Karampinos DC, Baum T, Nardo L, Alizai H, Yu H, Carballido-Gamio J, et al. Characterization of the regional distribution of skeletal muscle adipose tissue in type 2 diabetes using chemical shift-based water/fat separation. J Magn Reson Imaging 2012;35:899–907. https://doi.org/10.1002/jmri.23512.
[9] Samagh SP, Kramer EJ, Melkus G, Laron D, Bodendorfer BM, Natsuhara K, et al. MRI quantification of fatty infiltration and muscle atrophy in a mouse model of rotator cuff tears: mri quantification mouse rotator cuff tears. J Orthop Res 2013;31:421–6. https://doi.org/10.1002/jor.22233.
[10] Komolka K, Albrecht E, Wimmers K, Michal JJ, Maak S. Molecular Heterogeneities of Adipose Depots - potential effects on adipose-muscle cross-talk in humans, mice and farm animals. J Genom 2014;2:31–44. https://doi.org/10.7150/jgen.5260.
[11] Goodpaster BH, Kelley DE, Thaete FL, He J, Ross R. Skeletal muscle attenuation determined by computed tomography is associated with skeletal muscle lipid content. J Appl Physiol 2000;89:104–10. https://doi.org/10.1152/jappl.2000.89.1.104.
[12] Torriani M, Townsend E, Thomas BJ, Bredella MA, Ghomi RH, Tseng BS. Lower leg muscle involvement in Duchenne muscular dystrophy: an MR imaging and spectroscopy study. Skelet Radio 2012;41:437–45. https://doi.org/10.1007/s00256-011-1240-1.
[13] Elder CP, Apple DF, Bickel CS, Meyer RA, Dudley GA. Intramuscular fat and glucose tolerance after spinal cord injury – a cross-sectional study. Spinal Cord 2004;42:711–6. https://doi.org/10.1038/sj.sc.3101652.
[14] Gorgey AS, Dudley GA. Skeletal muscle atrophy and increased intramuscular fat after incomplete spinal cord injury. Spinal Cord 2007;45:304–9. https://doi.org/10.1038/sj.sc.3101968.
[15] Biltz NK, Meyer GA. A novel method for the quantification of fatty infiltration in skeletal muscle. Skelet Muscle 2017:7. https://doi.org/10.1186/s13395-016-0118-2.

[16] Liu W, Liu Y, Lai X, Kuang S. Intramuscular adipose is derived from a non-Pax3 lineage and required for efficient regeneration of skeletal muscles. Dev Biol 2012; 361:27–38. https://doi.org/10.1016/j.ydbio.2011.10.011.

[17] Minamoto VB, Suzuki KP, Bremner SN, Lieber RL, Ward SR. Dramatic changes in muscle contractile and structural properties after 2 botulinum toxin injections: muscle Dysfunction after 2 Botox ® Injections. Muscle Nerve 2015;52:649–57. https://doi.org/10.1002/mus.24576.

[18] Rowshan K, Hadley S, Pham K, Caiozzo V, Lee TQ, Gupta R. Development of fatty atrophy after neurologic and rotator cuff injuries in an animal model of rotator cuff pathology. J Bone Jt Surg 2010;92:2270–8. https://doi.org/10.2106/JBJS.I.00812.

[19] Kim HM, Galatz LM, Lim C, Havlioglu N, Thomopoulos S. The effect of tear size and nerve injury on rotator cuff muscle fatty degeneration in a rodent animal model. J Shoulder Elb Surg 2012;21:847–58. https://doi.org/10.1016/j.jse.2011.05.004.

[20] Gibbons MC, Singh A, Anakwenze O, Cheng T, Pomerantz M, Schenk S, et al. Histological evidence of muscle degeneration in advanced human rotator cuff disease. J Bone Jt Surg 2017;99:190–9. https://doi.org/10.2106/JBJS.16.00335.

[21] Valencia AP, Lai JK, Iyer SR, Mistretta KL, Spangenburg EE, Davis DL, et al. Fatty infiltration is a prognostic marker of muscle function after rotator cuff tear. Am J Sports Med 2018;46:2161–9. https://doi.org/10.1177/0363546518769267.

[22] Ogawa M, Lester R, Akima H, Gorgey A. Quantification of intermuscular and intramuscular adipose tissue using magnetic resonance imaging after neurodegenerative disorders. Neural Regen Res 2017;12:2100. https://doi.org/10.4103/1673-5374.221170.

[23] Hayot M. Skeletal muscle microbiopsy: a validation study of a minimally invasive technique. Eur Respir J 2005;25:431–40. https://doi.org/10.1183/09031936.05.00053404.

[24] Campagner A, Ciucci D, Svensson C-M, Figge MT, Cabitza F. Ground truthing from multi-rater labeling with three-way decision and possibility theory. Inf Sci 2021; 545:771–90. https://doi.org/10.1016/j.ins.2020.09.049.

[25] Svensson C-M, Hübler R, Figge MT. Automated classification of circulating tumor cells and the impact of interobsever variability on classifier training and performance. J Immunol Res 2015;2015:1–9. https://doi.org/10.1155/2015/573165.

[26] Rueden CT, Schindelin J, Hiner MC, DeZonia BE, Walter AE, Arena ET, et al. ImageJ2: imageJ for the next generation of scientific image data. BMC Bioinforma 2017;18:529. https://doi.org/10.1186/s12859-017-1934-z.

[27] Badrinarayanan V, Kendall A, Cipolla R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation. IEEE Trans Pattern Anal Mach Intell 2017. https://doi.org/10.1109/TPAMI.2016.2644615.

[28] Lloyd S. Least squares quantization in PCM. IEEE Trans Inf Theory 1982;28: 129–37. https://doi.org/10.1109/TIT.1982.1056489.

[29] Berg S, Kutra D, Kroeger T, Straehle CN, Kausler BX, Haubold C, et al. ilastik: interactive machine learning for (bio)image analysis. Nat Methods 2019;16: 1226–32. https://doi.org/10.1038/s41592-019-0582-9.

[30] Stringer C, Wang T, Michaelos M, Pachitariu M. Cellpose: a generalist algorithm for cellular segmentation. Nat Methods 2021;18:100–6. https://doi.org/10.1038/s41592-020-01018-x.

[31] Lundberg SM, Lee S-I. A Unified Approach to Interpreting Model Predictions. In: Guyon I, Luxburg UV, Bengio S, Wallach H, Fergus R, Vishwanathan S, et al., editors. Adv. Neural Inf. Process. Syst., vol. 30. Curran Associates, Inc.; 2017.

[32] Ribeiro MT, Singh S, Guestrin C. "Why Should I Trust You?": Explaining the Predictions of Any Classifier. Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min. San Francisco California USA: ACM,; 2016. p. 1135–44. https://doi.org/10.1145/2939672.2939778.

[33] Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. 2017 IEEE Int. Conf. Comput. Vis. ICCV. Venice: IEEE,; 2017. p. 618–26. https://doi.org/10.1109/ICCV.2017.74.

[34] Gerst R, Cseresnyés Z, Figge MT. JIPipe: visual batch processing for Image. J Nat Methods 2023. https://doi.org/10.1038/s41592-022-01744-4.

[35] P. Ehrlich. Z für Wiss Mikrosk und für Mikrosk Tech-Die Von Mir herrührende Hämatoxylinlösung 3 1986 133.

[36] Best DJ, Roberts DE. Algorithm AS 89: the upper tail probabilities of Spearman's Rho. Appl Stat 1975;24:377. https://doi.org/10.2307/2347111.

[37] Blumberg H. Book review: Grundzüge der Mengenlehre. Bull Am Math Soc 1920; 27:116–30. https://doi.org/10.1090/S0002-9904-1920-03378-1.

[38] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in python. J Mach Learn Res 2011;12:2825–30.

[39] Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. J Big Data 2019;6:60. https://doi.org/10.1186/s40537-019-0197-0.

[40] Ridler TW, Calvard S. Picture thresholding using an iterative selection method. IEEE Trans Syst Man Cyber 1978;8:630–2. https://doi.org/10.1109/TSMC.1978.4310039.

[41] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. arXiv 2015:234–41. https://doi.org/10.1007/978-3-319-24574-4_28.

[42] Cohen J. A power primer. Psychol Bull 1992;112:155–9. https://doi.org/10.1037/0033-2909.112.1.155.

[43] Moor K, Diard M, Sellin ME, Felmy B, Wotzka SY, Toska A, et al. High-avidity IgA protects the intestine by enchaining growing bacteria. Nature 2017;544:498–502. https://doi.org/10.1038/nature22058.

[44] Alahmari SS, Goldgof D, Hall LO, Mouton PR. A review of nuclei detection and segmentation on microscopy images using deep learning with applications to unbiased stereology counting. IEEE Trans Neural Netw Learn Syst 2022:1–20. https://doi.org/10.1109/TNNLS.2022.3213407.

[45] Ravindran S. Five ways deep learning has transformed image analysis. Nature 2022;609:864–6. https://doi.org/10.1038/d41586-022-02964-6.

[46] A. Kirillov E. Mintun N. Ravi H. Mao C. Rolland L. Gustafson et al. Segm Anything 2023 doi: 10.48550/ARXIV.2304.02643.

[47] Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, et al. Transformers for Image Recognition at Scale. arXiv 2020. https://doi.org/10.48550/ARXIV.2010.11929.

[48] Malmgren LT, Gacek RR. Histochemical characteristics of muscle fiber types in the posterior cricoarytenoid muscle. Ann Otol Rhinol Laryngol 1981;90:423–9. https://doi.org/10.1177/000348948109000503.

[49] Brøndbo K, Dahl HA, Teig E, Gujord KM. The Human Posterior Cricoarytenoid (PCA) Muscle and Diaphragm: a histochemical comparison as a basis for reinnervation attempts. Acta Otolaryngol (Stock) 1986;102:474–81. https://doi.org/10.3109/00016488609119433.

[50] Happak W, Zrunek M, Pechmann U, Streinzer W. Comparative histochemistry of human and sheep Laryngeal Muscles. Acta Otolaryngol (Stock) 1989;107:283–8. https://doi.org/10.3109/00016488909127510.

[51] Ro K, Kim JY, Park H, Cho BH, Kim IY, Shim SB, et al. Deep-learning framework and computer assisted fatty infiltration analysis for the supraspinatus muscle in MRI. Sci Rep 2021;11:15065. https://doi.org/10.1038/s41598-021-93026-w.

[52] Weber KA, Smith AC, Wasielewski M, Eghtesad K, Upadhyayula PA, Wintermark M, et al. Deep learning convolutional neural networks for the automatic quantification of muscle fat infiltration following whiplash injury. Sci Rep 2019;9:7973. https://doi.org/10.1038/s41598-019-44416-8.

[53] Weber KA, Abbott R, Bojilov V, Smith AC, Wasielewski M, Hastie TJ, et al. Multi-muscle deep learning segmentation to automate the quantification of muscle fat infiltration in cervical spine conditions. Sci Rep 2021;11:16567. https://doi.org/10.1038/s41598-021-95972-x.

[54] Estrada S, Lu R, Conjeti S, Orozco-Ruiz X, Panos-Willuhn J, Breteler MMB, et al. FatSegNet: a fully automated deep learning pipeline for adipose tissue segmentation on abdominal dixon MRI. Magn Reson Med 2020;83:1471–83. https://doi.org/10.1002/mrm.28022.

[55] Cseresnyes Z, Hassan MIA, Dahse H-M, Voigt K, Figge MT. Quantitative impact of cell membrane fluorescence labeling on phagocytosis measurements in confrontation assays. Front Microbiol 2020;11:1193. https://doi.org/10.3389/fmicb.2020.01193.

[56] Joyce NC, Oskarsson B, Jin L-W. Muscle biopsy evaluation in neuromuscular disorders. Phys Med Rehabil Clin N Am 2012;23:609–31. https://doi.org/10.1016/j.pmr.2012.06.006.

[57] Walters J, Baborie A. Muscle biopsy: what and why and when. Pr Neurol 2020;20: 385–95. https://doi.org/10.1136/practneurol-2019-002465.

[58] Goutallier D, Postel JM, Bernageau J, Lavau L, Voisin MC. Fatty muscle degeneration in cuff ruptures. Pre- and postoperative evaluation by CT scan. Clin Orthop 1994:78–83. https://doi.org/PMID: 8020238.

[59] Nakagaki K, Ozaki J, Tomita Y, Tamai S. Fatty degeneration in the supraspinatus muscle after rotator cuff tear. J Shoulder Elb Surg 1996;5:194–200. https://doi.org/10.1016/S1058-2746(05)80005-9.

[60] Visser M, Goodpaster BH, Kritchevsky SB, Newman AB, Nevitt M, Rubin SM, et al. Muscle Mass, Muscle Strength, and muscle fat infiltration as predictors of incident mobility limitations in well-functioning older persons. J Gerontol A Biol Sci Med Sci 2005;60:324–33. https://doi.org/10.1093/gerona/60.3.324.

[61] Leroy-Willig A, Willig TN, Henry-Feugeas MC, Frouin V, Marinier E, Boulier A, et al. Body composition determined with MR in patients with Duchenne muscular dystrophy, spinal muscular atrophy, and normal subjects. Magn Reson Imaging 1997;15:737–44. https://doi.org/10.1016/S0730-725X(97)00046-5.

[62] Nassoro DD, Torres L, Marando R, Mboma L, Mushi S, Habakkuk Mwakyula I. A child with duchenne muscular dystrophy: a case report of a rare diagnosis among Africans. Clin Case Rep 2020;8:2654–60. https://doi.org/10.1002/ccr3.3254.

[63] Goodpaster BH, Wolf D. Skeletal muscle lipid accumulation in obesity, insulin resistance, and type 2 diabetes. Pedia Diabetes 2004;5:219–26. https://doi.org/10.1111/j.1399-543X.2004.00071.x.

[64] Cui H, Liu L, Liu X, Wang Y, Luo N, Tan X, et al. A selected population study reveals the biochemical mechanism of intramuscular fat deposition in chicken meat. J Anim Sci Biotechnol 2022;13:54. https://doi.org/10.1186/s40104-022-00705-3.

[65] Gotoh T, Nishimura T, Kuchida K, Mannen H. The Japanese Wagyu beef industry: current situation and future prospects - a review. Asian-Austral J Anim Sci 2018; 31:933–50. https://doi.org/10.5713/ajas.18.0333.

[66] Rivenson Y, Wang H, Wei Z, de Haan K, Zhang Y, Wu Y, et al. Virtual histological staining of unlabelled tissue-autofluorescence images via deep learning. Nat Biomed Eng 2019;3:466–77. https://doi.org/10.1038/s41551-019-0362-y.

[67] Greenwald NF, Miller G, Moen E, Kong A, Kagel A, Dougherty T, et al. Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning. Nat Biotechnol 2022;40:555–65. https://doi.org/10.1038/s41587-021-01094-0.

[68] Lauer J, Zhou M, Ye S, Menegas W, Schneider S, Nath T, et al. Multi-animal pose estimation, identification and tracking with DeepLabCut. Nat Methods 2022;19: 496–504. https://doi.org/10.1038/s41592-022-01443-0.