WormCNN

Subjects: Biotechnology & Applied Microbiology
Contributor: Yan Pan

WormCNN is a specialized deep learning model designed for the analysis of images of the nematode *Caenorhabditis elegans* (*C. elegans*). This model utilizes convolutional neural network (CNN) architecture to classify and analyze the morphological features of *C. elegans*, providing valuable insights into various biological studies.

C. elegans CNN Deep Learning

1. Introduction

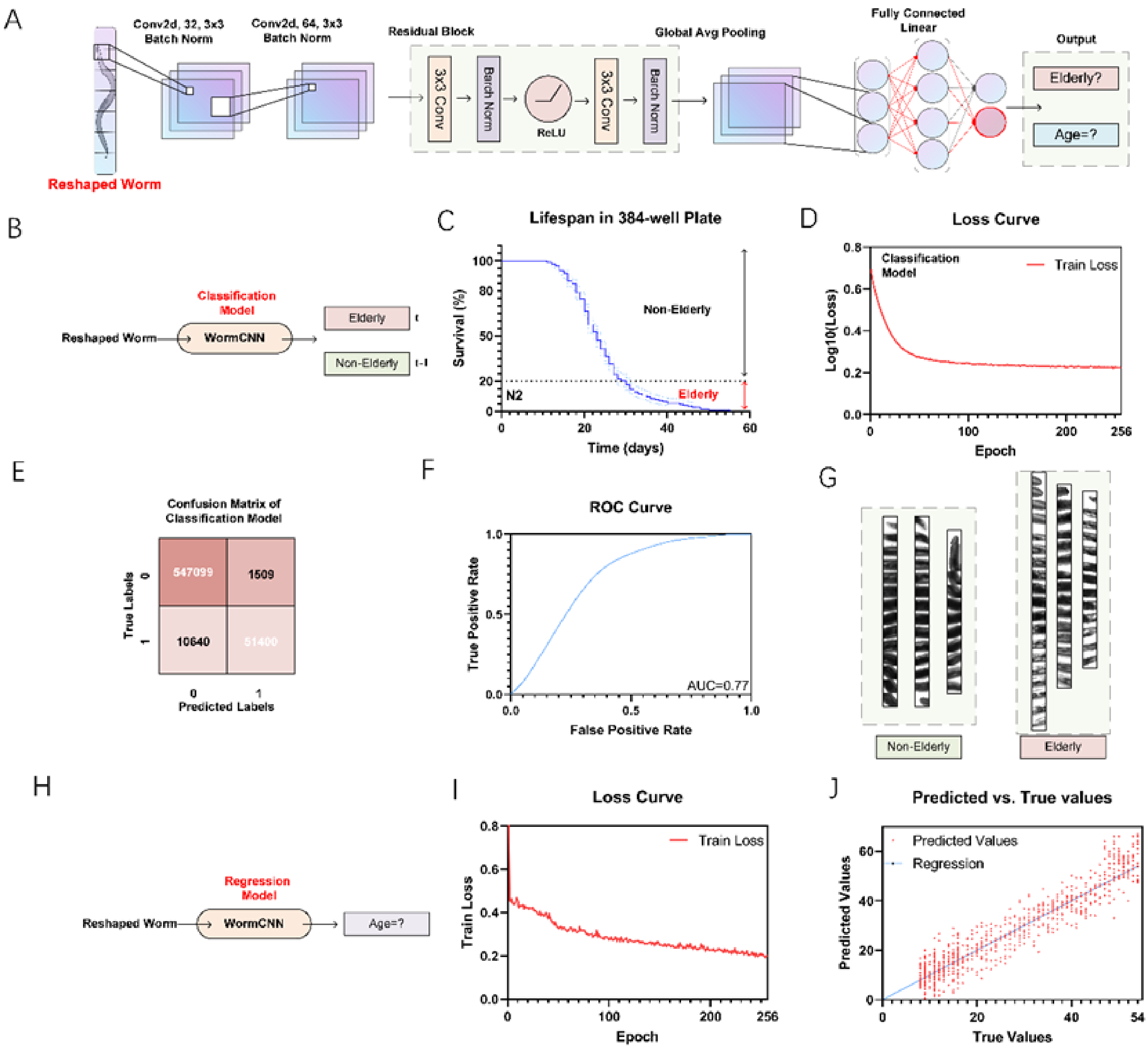
Caenorhabditis elegans is a small free-living nematode widely used as a model organism in biological and genetic research due to its simple anatomy, short lifecycle, and well-annotated genome. Analyzing *C. elegans* images is crucial for studies on aging, development, and the effects of genetic mutations. Traditional image analysis methods are time-consuming and require human expertise. WormCNN addresses these challenges by offering an automated and efficient deep learning-based image analysis solution.

WormCNN is a convolutional neural network (CNN) model designed for age classification and regression analysis of *C. elegans* images . The authors introduce a method called SingleWorm, where individual worms are cultured in liquid in 384-well plates, and daily images of surviving worms are captured as a dataset.

After training the model, the predicted biological age is compared with reference data based on lifespan indicators in liquid culture. This allows for calculating the “Healthy Aging Index” (HAI), which quantitatively assesses the aging health of worms. The model also enables binary classification of worms as "young" or "old."

WormCNN introduces a method for linearizing worm images.

Similar to tokenization in natural language processing, the WormCNN method divides worm images into multiple "tokens" based on the overall direction of the worm. These tokens are then recombined to create a linearized version of the worm image. This preprocessing step is crucial in reducing variations caused by distortions, bends, or other morphological changes commonly seen in worm movement. By linearizing the images, the model can focus on detecting biological feature changes without being distracted by pose differences. This standardization improves the overall accuracy of classification and regression tasks.



2. Preprocessing Details

Image capture and data collection: Worm images are captured daily using the SingleWorm method in a controlled environment, where worms are cultured individually in 384-well plates. This method minimizes interactions between worms and ensures that the images collected accurately represent the morphological state of each individual without interference from other worms. Images are typically captured in grayscale, simplifying processing and reducing computational load.

Skeletonization and feature extraction: During preprocessing, each worm's image is converted into a skeleton structure. This process involves detecting the worm's body and simplifying it into a skeleton line, preserving key structural features like body curvature while discarding irrelevant details. Skeletonization reduces the complexity of the input data, making it easier for the model to focus on relevant morphological features.

Resampling skeleton coordinates: The skeleton coordinates are resampled to generate dense skeleton points, providing a better representation of the worm's structure. This ensures data consistency, with each image represented by a fixed number of skeleton points. These points are used to facilitate the linearization process.

Linearization of worm images: The resampled skeleton points are then used to linearize the worm. The worm's body is straightened along its overall direction, dividing the image into small segments or "tokens," which are then reassembled to mimic a straightened worm. This step reduces variations caused by different poses, such as curling or bending, enabling the model to analyze the worms in a standardized form.

Morphological feature standardization: By converting each worm into a standardized linear image, WormCNN ensures that morphological features (such as body thickness, length, and other age-related changes) remain consistent across all images. This standardized representation allows the model to focus on subtle biological feature changes related to aging or health status without being distracted by irrelevant pose variations.

3. WormCNN Architecture

The WormCNN model is built on a convolutional neural network framework, which is well-suited for image recognition tasks. The architecture consists of multiple layers designed to automatically learn and extract relevant features from *C. elegans* images. The workflow of the model can be summarized as follows:

Input layer: The model takes in preprocessed images of *C. elegans*, typically in grayscale, for simplicity and reduced computational complexity.

Convolutional layers: These layers are responsible for feature extraction. WormCNN uses multiple sets of filters to detect various features in the images, such as edges, curves, and textures.

Activation functions: Non-linear activation functions, such as the rectified linear unit (ReLU), introduce non-linearity into the model, allowing it to learn more complex patterns.

Pooling layers: These layers reduce the spatial dimensions of the feature maps, decreasing the number of parameters and computations in the network.

Fully connected layers: The final layers of the network integrate the extracted features and make classification decisions.

Output layer: WormCNN typically uses a sigmoid activation function in the output layer for binary classification tasks, such as distinguishing between young and old worms.

4. Applications

WormCNN has a wide range of applications in biological research:

- Aging research: By classifying *C. elegans* by age based on their morphology, WormCNN can contribute to studies on aging and the effects of various interventions on lifespan.
- Developmental biology: The model can be used to track developmental stages and identify the effects of any abnormalities or genetic mutations.
- Genetic research: In studies involving gene editing or knockouts, WormCNN can assist in phenotype analysis by examining the worms' physical traits.
- High-throughput screening: In laboratories requiring large-scale image analysis, WormCNN can automate the process and improve data analysis efficiency.

5. Conclusion

WormCNN represents a novel approach in computational biology, providing a tool for automated image analysis of *C. elegans*. Its ability to accurately classify and analyze images of these worms has the potential to accelerate research in various biological fields.

References

1. Yan Pan; Zhihang Huang; Hongxia Cai; Zhiru Li; Jingyuan Zhu; Dan Wu; Wentao Xu; Hexiang Qiu; Nan Zhang; Guojun Li; Shan Gao; Bo Xian; WormCNN-Assisted Establishment and Analysis of Glycation Stress Models in C. elegans: Insights into Disease and Healthy Aging. *Int. J. Mol. Sci.*. **2024**, *25*, 9675.