

RESEARCH

Open Access



Web-based diagnostic platform for microorganism-induced deterioration on paper-based cultural relics with iterative training from human feedback

Chenshu Liu^{1*†}, Songbin Ben^{2,3*†}, Chongwen Liu⁴, Xianchao Li², Qingxia Meng², Yilin Hao², Qian Jiao² and Pinyi Yang⁵

Abstract

Purpose Paper-based artifacts hold significant cultural and social values. However, paper is intrinsically fragile to microorganisms, such as mold, due to its cellulose composition, which can serve as a microorganisms' nutrient source. Mold not only can damage papers' structural integrity and pose significant challenges to conservation works but also may subject individuals attending the contaminated artifacts to health risks. Current approaches for strain identification usually require extensive training, prolonged time for analysis, expensive operation costs, and higher risks of secondary damage due to sampling. Thus, in current conservation practices with mold-contaminated artifacts, little pre-screening or strain identification was performed before mold removal, and the cleaning techniques are usually broad-spectrum rather than strain-specific. With deep learning showing promising applications across various domains, this study investigated the feasibility of using a convolutional neural network (CNN) for fast in-situ recognition and classification of mold on paper.

Methods Molds were first non-invasively sampled from ancient Xuan Paper-based Chinese books from the Qing and Ming dynasties. Strains were identified using molecular biology methods and the four most prevalent strains were inoculated on Xuan paper to create mockups for image collection. Microscopic images of the molds as well as their stains situated on paper were collected using a compound microscope and commercial microscope lens for cell phone cameras, which were then used for training CNN models with a transfer learning scheme to perform the classification of mold. To enable involvement and contribution from the research community, a web interface that actuates the process while providing interactive features for users to learn about the information of the classified strain was constructed. Moreover, a feedback functionality in the web interface was embedded for catching potential classification errors, adding additional training images, or introducing new strains, all to refine the generalizability and robustness of the model.

Results & Conclusion In the study, we have constructed a suite of high-confidence classification CNN models for the diagnostic process for mold contamination in conservation. At the same time, a web interface was constructed

[†]Chenshu Liu and Songbin Ben have contributed equally to this work.

*Correspondence:

Chenshu Liu
chenshu0615@g.ucla.edu
Songbin Ben
bensongbin@lnu.edu.cn

Full list of author information is available at the end of the article



that allows recurrently refining the model with human feedback through engaging the research community. Overall, the proposed framework opens new avenues for effective and timely identification of mold, thus enabling proactive and targeted mold remediation strategies in conservation.

Keywords Paper-based cultural relics, Conservation, Computer vision, Deep learning, Strain classification, Web application, Iterative training, Human feedback

Introduction

Paper has been one of the main mediums of communication and knowledge representation throughout history and across the world. However, due to its cellulose composition, paper is susceptible to various physical, chemical, and biological agents present in the environment, particularly to microorganisms such as molds [1]. Aged papers are especially prone due to spontaneous and environment-induced acidification, which creates a hospitable environment for microorganisms to grow and reproduce [2]. The byproducts of microorganisms' growth and metabolism contain enzymes and acidic compounds that will disintegrate the fibrous cellulose structure of the paper, resulting in both aesthetic and mechanical deterioration to the paper material itself as well as the information it carries [3]. Furthermore, given the omnipresence of microorganisms [4], there are constant threats of contamination to paper-based materials. The contamination process is usually undetectable by the naked eye, making it difficult to intervene at an early stage [5]. When the contamination is visibly apparent, a substantial colony composed of microorganisms has already been engraved in the cellulose structure of the paper, causing irreversible deterioration, making the time window for effective remedial measures highly limited. Thus, the impacts of microorganisms on books, documents, and other paper-based materials may result in inestimable cultural losses [1, 6] and may induce huge financial burdens [7, 8]. Therefore, effective detection and treatment of bio-contamination due to microorganisms' contamination are crucial for the preservation of unique artifacts for humanity and future generations [9].

Current mold remediation practices in conservation are generally wide-spectrum. A general workflow consists of mechanical cleaning of microorganism bodies through brushing and vacuum cleaning and sometimes coupled with chemical or physical treatment to further eradicate the microorganism spores [10]. Chemical methods involve fumigation using membrane-active microbicides, such as alcohol, salicylanilides, and quaternary ammonium salts. These microbicides coat the cell wall and then cause damage to the structure of the mold bodies. Alternatively, electrophilically active microbicides such as aldehydes and organometallic compounds act through electrophilic addition or substitution to

cause enzyme inactivation [1]. These wide-spectrum biocides, however, could induce more rapid mutations in the microorganism strain, and elevate environmental risks and public health concerns [11]. On the other hand, physical techniques involve dehydration, radiation, and deoxygenation that perturb the hospitable environmental conditions that are essential for molds to survive and further reproduce [1]. However, the extreme environmental conditions applied in the physical methods could subject the delicate ancient paper material to secondary damage. The development of species-targeted approaches has the potential to minimize the risk of secondary destruction and aligns with the conservation principle of "minimum intervention" [12]. For example, species-targeted biocides could be designed based on different biocidal resistances observed in different microorganism species [13]. After all, successful implementation of such treatments relies heavily on fast and accessible microorganism identification methods.

Additionally, the identification of mold species has the potential to resolve practical and safety problems in conservation work settings. First, different consortia of microorganisms display varying mechanisms of destruction [6], making it challenging for conservators to provide an accurate condition report and propose the optimal treatment options without knowing the contaminating species. However, if the strain of the mold is known when diagnosing the material, the best approach to conserve can be devised accordingly. More importantly, according to the Centers for Disease Control and Prevention (CDC), mold is associated with many health complications. People who have access to or frequently attend to the contaminated materials are subjected to potential health risks given the presence of mold [6, 14]. Molds can invade living tissues and cause diseases or allergies, such as allergic skin, respiratory symptoms, or gastrointestinal disorders [6, 15]. The pathogenicity, transmission route, and infection severity are also distinctive with various species. As a result, to ensure efficiency and safety in the treatment of microbial contamination, species of the mold should be carefully identified and evaluated [3].

The identification of mold species before conservation practices is not common in the current conservation field. This is primarily because the mainstream identification techniques, namely culture-based and molecular

biological methods, are not widely accessible from the conservators' point of view [15]. Culture-based techniques rely heavily on culturing the microorganisms in a qualified laboratory environment, which is time-consuming in terms of both making connections with specialized labs and conducting the analysis. Meanwhile, molecular biology techniques that directly analyze the sequence of the target organisms, even though has high authenticity and accuracy, take a prolonged period to generate results [8, 16]. For example, pyrosequencing, which is based on the luminometric detection of pyrophosphate can achieve an accuracy of 99% but a single run takes 24 h [17]. Illumina, a sequencing method based on a reversible dye-terminator and engineered polymerase can achieve single-shot accuracy of 99.9%, however, the operation time is 1–11 days [18]. Another significant drawback of sequencing techniques is that, to obtain sufficient amounts of sample, they require direct sampling from the object being conserved, which is considered invasive and would induce secondary damage, especially to delicate historical material [3]. In summary, fast and non-invasive tools are still absent for conservators and other practitioners in everyday work.

AI-enabled applications are gradually integrating into various aspects of our lives. Benefiting from the quick advancements in artificial neural networks called deep neural networks (DNN), more and more challenging tasks that once required profound human expertise can now be achieved by computers [16, 19]. In particular, computer vision (CV), a branch of DNN application that enables computers to identify and understand objects and people in images and videos [19], is gaining momentum in assisting with expert-level image and video-related tasks and can capture intricate features that may not be apparent to the naked eye [20]. For example, CV has shown huge potential in the field of medicine to enhance healthcare workflows [20]. A team of experts in Poland applied DNN and a bag-of-words approach to classify microscopic images of various fungal species. The approach significantly reduced the diagnostic time from 4 to 10 days down to less than 3 days. This not only allows faster decisions regarding antifungal treatments to be prescribed to patients who had a fungal infection, resulting in shorter recovery time but also reduces the cost of diagnosis due to the replacement of biochemical testing procedures [21]. There have also been attempts to integrate DL techniques into conservation practices. For example, Hatir et al. proposed to use deep neural networks to assist with the task of classifying the weathering condition for historical stone monuments which can help in conservation and restoration practices [22]. Another team in China proposed using FSNets to automatically recognize and count fungal spores in microscopic images

to help monitor grain storage and detect signs of spoiling and fungi contamination [23]. Given the nearly real-time feedback and minimal invasiveness, CV is an ideal candidate for diagnosing delicate cultural heritage like paper. Thus, this study investigated the feasibility of a CV-based diagnostic tool for identifying mold strains on paper-based relics non-invasively to assist conservation practices. Convolutional Neural Network (CNN) architectures, a type of mainstream CV algorithm, were experimented with to classify the strains with microscopic images. At the same time, the study provides a framework to allow the tool to grow in tandem with the growing knowledge in the field of conservation.

Methods

Sample preparation

Microscopic images of the visually apparent stained regions on paper were collected for modeling “Modeling” section. The sample preparation process involves extracting the molds from ancient artifacts, creating contaminated mockups for the ancient artifacts, and collecting microscopic images for model training (Fig. 1).

Mold acquisition

Microorganisms were sampled and revived from four ancient archives with minimal intervention, including “Sequel of Comprehensive Reflections to Aid in Governance” from the Ming Dynasty, “Supervised Copy of I Ching” from the Qing Dynasty, “True Interpretation of Journey to the West” from Qing dynasty, and “Veritable Records of the Qing Dynasty” from Qing dynasty (Fig. 2B), that were stored in the department of rare books and special collection at Liaoning University Library. Over 10 different strains of microorganism were resuscitated from the books, and four strains of molds that were most commonly found to contaminate paper-based materials and pose more serious threat were selected for this project: *Aspergillus niger* (*A. Niger*), *Aspergillus ochraceus* (*A. Oschraceus*), *Cladosporium*, and *Paenibacillus polymyxa* (*P. Polymyxa*) [6, 24–27]. *A. Niger* is one of the most commonly mentioned molds in literature. Besides producing cellulolytic enzymes that can digest cellulose, *A. Niger* can also secrete a wide spectrum of biological enzymes, including amylase, pectinase, and protease, whose by-products, can increase the overall acidity in the paper, and then induce severe loss of papers' folding endurance in the long-term [25, 28]. *Cladosporium* can also secrete cellulolytic enzymes. Even though its activity is not as high as that of *A. Niger*, *Cladosporium* is capable of secreting pigments with high concealment over the original material [29–31]. *P. Polymyxa* is a type of bacteria that is not only potent in protein hydrolysis and cellulose degradation but also tends

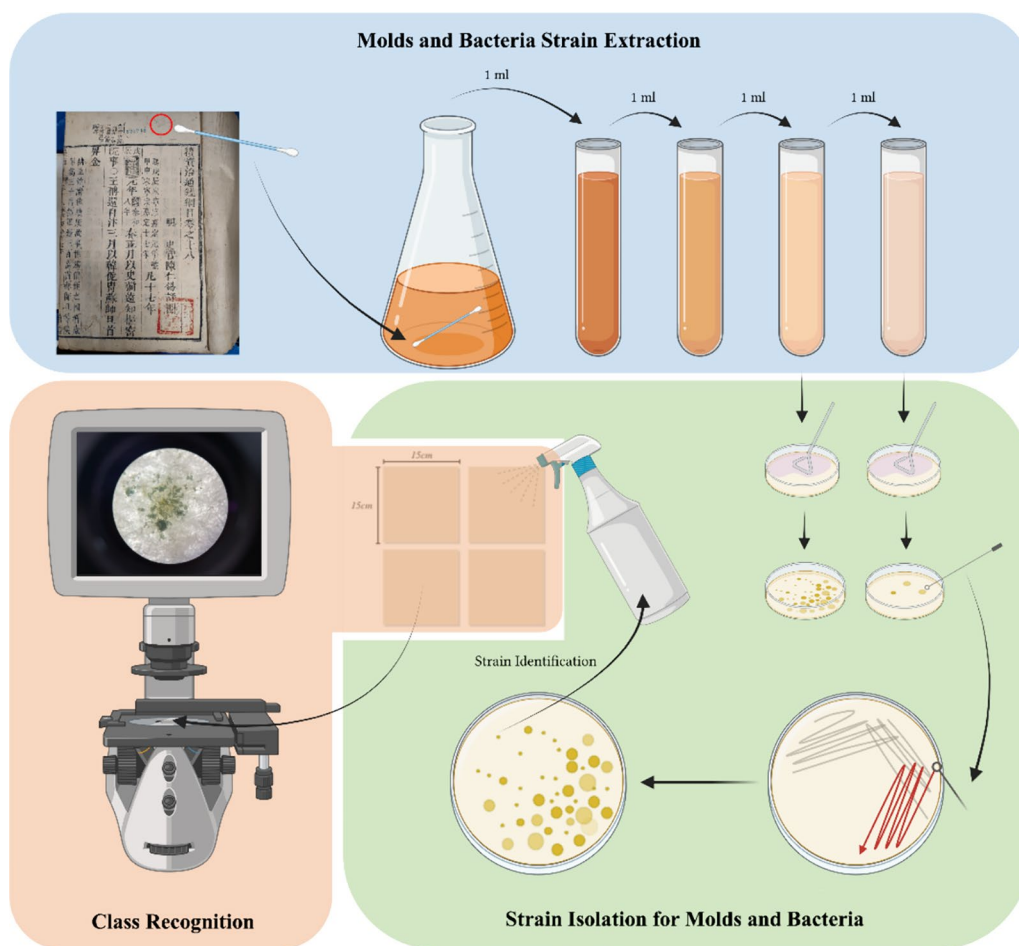


Fig. 1 Sample acquisition pipeline

to secrete humus compounds consisting of polysaccharide, lipid, protein, and nucleic acid, which cause adhesion between papers [32, 33]. *A. Oschraceus* was selected for its potency in producing spores and causing large-scale air contamination that may induce asthma and even cancer for people who have been in contact with the contaminated materials [34]. The following step-by-step procedures outline the strain extraction process:

1. Visually identify regions of paper that show significant colonization of microorganism contamination from ancient books.
2. Gently swab the discolored region to collect microorganism strains and place the swab, together with liquid beef extract peptone medium (BPM), in a conical flask and culture the suspension in a lab shaker at 37 °C for 125r/min for 12–72 h.
3. Use dilution plating to transfer bacteria and fungus from liquid-state medium to solid-state BPM and potato dextrose agar medium (PDA) respectively and cultivate for 1–3 days at 28 °C for initial isolation and purification.
4. Select an isolated colony for streak plating and cultivate for 2–7 days at 28 °C for secondary isolation and purification.
5. Store the isolated strains using slant streaking at 4°C for later use.
6. Use microorganism colony characteristics for preliminary strain classification and then use 16 s rDNA, and 18 s/ITS rDNA for strain determination for bacteria and fungus respectively.
7. Select the aforementioned four microorganism strains and suspend them in a liquid medium. Place in lab shaker at 28 °C for 125r/min to further proliferate.
8. Centrifuge the suspension when reaches the logarithmic growth phase, collect condensation (microorganism bodies), and re-suspend in sterile water. Save the solution containing the suspended microorganisms for creating mockup samples.

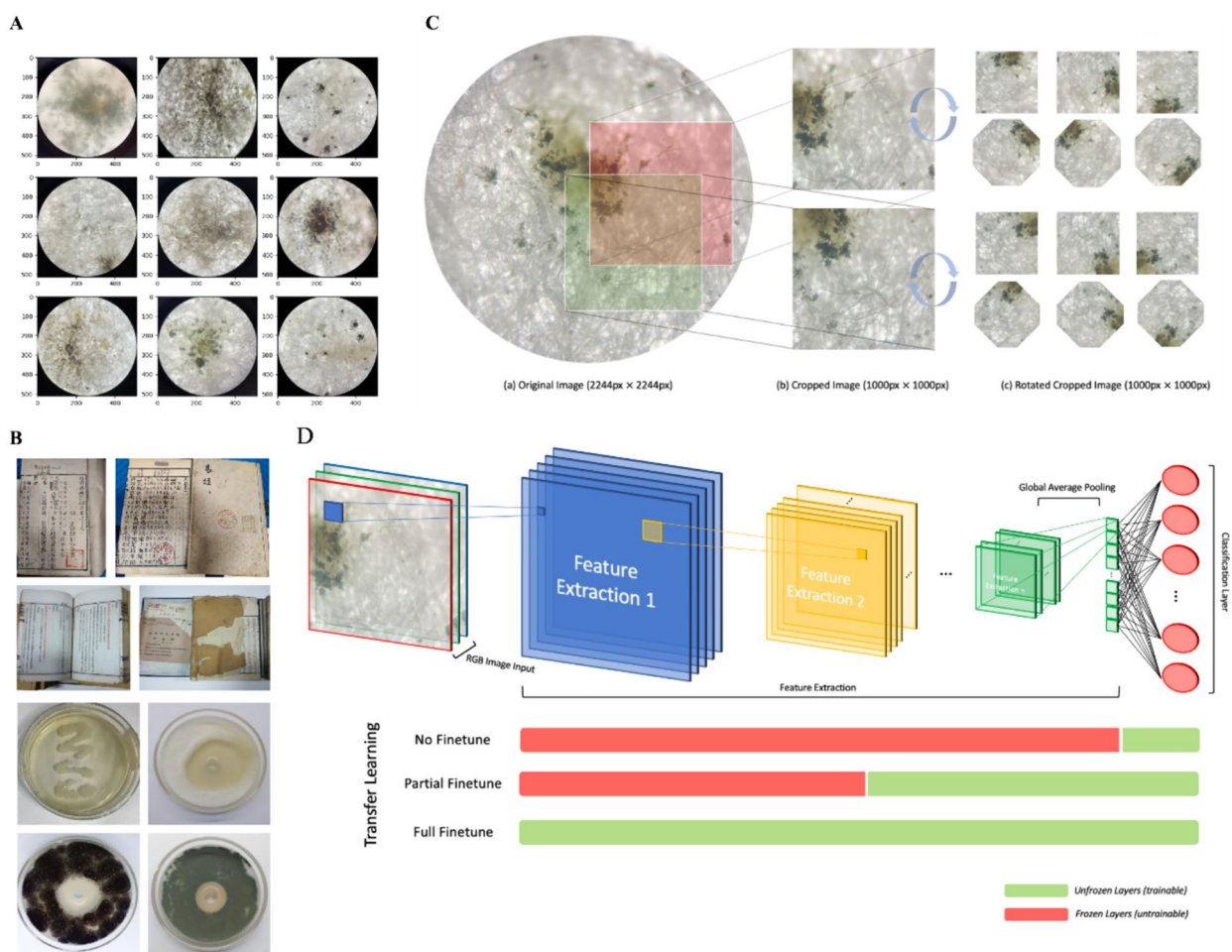


Fig. 2 Image processing and modeling pipeline. **A** Microscopic images of the strains being classified. **B** Ancient paper artifacts were used to retrieve the mold strains, and strain growth in the culturing environment. **C** Image augmentation pipeline with random cropping and rotation. **D** Transfer learning modeling architecture with different partial training paradigms

9. Bio-waste disposal: The residual material, including all apparatus that came in contact with the microorganism such as the contaminated paper, underwent neutralization procedures as enforced by the health administration agency.

Mockup construction

Because the ancient documents that the molds were sampled from are fragile due to natural aging and are off the limit for frequent access, mockups of mold-contaminated papers were constructed. The microorganism suspension solutions, obtained in “Mold acquisition” section, are evenly sprayed onto pre-cut 15 cm x 15 cm papers using a fog spray, with three duplicate samples for each strain. The sprayed papers are then sealed in Ziplock bags to avoid contamination and inoculated for 7–10 days in an incubator at 28 °C. As a way to mimic the growth

environment of microorganisms on paper-based relics, routinely spray sterile water onto contaminated papers every 12 h to facilitate the growth and reproduction of the microorganisms.

Image collection

Microscopic images of the mold-contaminated papers, obtained from step 2.1.2, were collected for the four strains of microorganisms with three different magnifications, including 10x and 40x from a standard laboratory compound microscope, as well as one from a commercially available phone-attached microscope lens that was bought online (Fig. 2A). The images obtained with the commercial microscope lens were included to enhance the accessibility and portability of the model. Costing less than \$40 on Amazon, the lens would allow people to capture microscopic features similar to the ones captured by the standard laboratory

microscope. Additionally, microscopic images without any microorganisms under the field of view were added to the image collection to serve as blank controls, allowing the CV algorithm to only learn about the representative features of the mold instead of the paper cellulose background.

Modeling

The modeling process contains the training phase and the testing phase. The training phase involves training a CNN model to identify mold species. Images are first labeled with the corresponding classes, in this case, the strains, and then undergo data preprocessing to obtain the training data before being fed into the training process. The model gradually improves itself over training epochs until it achieves optimal results for the set of evaluation criteria. The trained models are further validated using testing data, a completely new set of images that the model had not seen before, to evaluate the robustness of the classification model (Fig. 3).

The process was carried out in the cloud-based Jupyter Notebook environment provided by Google Colab, using NVIDIA A100 GPU. The deep neural network architecture was constructed using the TensorFlow API version 2.11.0.

Data preprocessing

The image dataset was divided into training and testing sets before further preprocessing. This prevents the risk of testing images inadvertently being included in the training set, often referred to as data leakage, which prevents the model from memorizing specific images and causing a pseudo-confident model. Further, given the limited number of images that can be generated from ancient books with highly restricted access, data augmentation is used to augment the number of images to reduce the risk of overfitting. Overfitting refers to the situation that, when the number of training samples is limited, the models will become overly familiar with the training data that they almost remember certain features but cannot generalize well to features that it was not exposed to before. Traditional data augmentation techniques for image data include rotating the images within a range of angles, translating or flipping the images in various directions, and cropping the images [35]. In this study, data augmentation was achieved through random cropping and rotation (Fig. 2C). The crop was generated by first initializing the x and y coordinates of a random starting point (i.e. the upper-left corner coordinate) and pinpointing the other three corner coordinates to create 1000pixels×1000pixels images. The randomized cropping procedure was repeated on each image for 50 times.

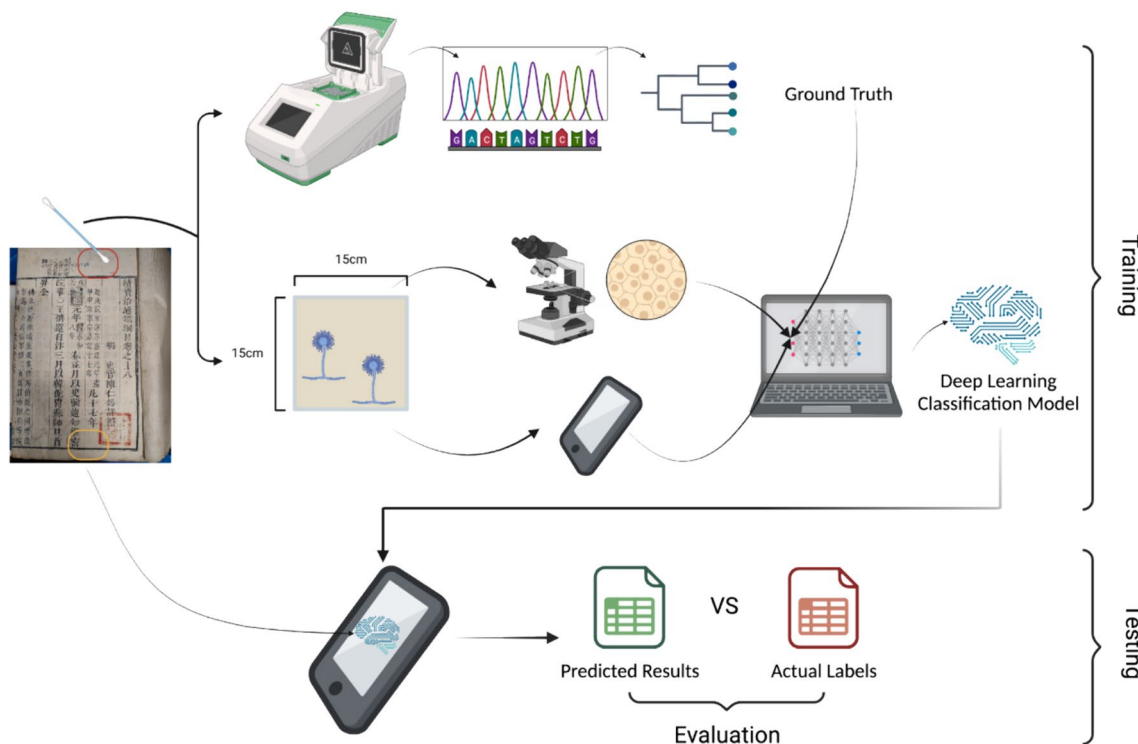


Fig. 3 Modeling Pipeline

Then, the cropped sections were rotated in angles of $n \times \frac{\pi}{8}$, where $0 \leq n < 8$, to further augment the number of images. In the end, images that do not contain sufficient representative features were manually discarded from the image dataset to avoid false positives in the learning process.

Modeling scheme

In this study, a transfer learning scheme was used for training the classification model. To further prevent overfitting, transfer learning offers a solution by leveraging the knowledge, represented as weights in the model, gained from off-the-shelf pre-trained models that were trained on well-rounded datasets, such as ImageNet, a large natural image database containing over 14 million images of 20,000 types of objects. Through transfer learning, the development pipeline is simplified by avoiding training a new model all from scratch [36]. The modified model is then trained on data from the target domain to further tune the model weights, through a process called fine-tuning, to adapt to more niche tasks [35, 37–40].

The train/freeze trade-off was considered when fine-tuning the transfer learning scheme used in this study. The initial weights of the pre-trained models enable the model to perform well on the pretraining dataset (i.e. ImageNet) so that, during the later fine-tuning on the target data, knowledge learned about general graphical features can be transferred, rather than learning from scratch. In practice, the transferred knowledge was retained at various levels by freezing different numbers of layers in the model. Freezing more layers in the pre-trained model can reduce the amount of computation needed. However, when the pretraining dataset is largely different from the target dataset, it would make the model underfitting on the target features. On the other hand, unfreezing more layers in the pre-trained model can increase the fitting performance on the target set. However, unfreezing too many would be equivalent to training the model from scratch, which demands longer training time and more computational resources.

In this study, three types of finetuning schemes were experimented with, considering the train/freeze trade-off, namely full finetuning, partial finetuning, and no fine-tuning (Fig. 2D). No finetuning refers to setting the feature extraction block that was transferred from the pre-trained models to untrainable, and only the classification layer is set to be trainable so that the architecture can adapt to the new classification task. Partial finetuning refers to partially unfreezing the feature extraction block to allow some layers to be trainable with the new set of data while the rest are not trainable. In particular, shallow layers of the feature extraction layer that capture higher-level information are set to be trainable to better

learn the features that are unique to the new image set, while deeper layers that capture lower-level information such as edges, lines, and colors, are set to be non-trainable because those are characteristics shared by nearly all objects. Finally, the full finetune refers to setting the overall architecture, including the transferred feature extraction layers as well as the classification layer, to be trainable by the new image data.

Model architecture

Convolutional Neural Networks (CNN) is one of the mainstream algorithms powered by DL [41], widely explored in numerous CV application scenarios, such as medicine [42], robotics [43], etc. CNN models are achieved with supervised learning, where a large number of images are presented with labels to the model, then the model, which is composed of a series of layers to detect different features of the input images, will gradually learn the representative features corresponding to each label.

Visual Geometry Group (VGG) and Residual Network (ResNet) architectures, both widely applied CNN architectures [41, 44], were investigated in this study because they have shown consistent performance in microscopic image classification through training with large amounts of data or transfer learning [16, 39].

The feature extraction layers of the two pre-trained networks, VGG16 and ResNet50, are transferred for the new classification task, as explained in “**Modeling Scheme**” section. A Global Average Pooling (GAP) layer, proposed as an alternative for fully connected layers in classical CNNs [45], is added after the output of the feature extraction to take the average of each of the feature maps from the last layer of the feature extraction block and flatten it into a one-dimensional vector. GAP layers can efficiently reduce the spatial dimensions of three-dimensional feature maps by downsampling the entire feature map to a single value [46]. The one-dimensional output from the GAP layer is then passed into a fully connected layer with five nodes, each representing a class label of the classification task (Fig. 2D).

Modeling experiment

A dynamic adjustment of the number of training epochs and learning rate was experimented in the study. Early stopping is implemented with pre-defined stopping criteria during the training process to avoid overfitting. In the study, parameters including categorical cross-entropy loss, accuracy, recall, precision, and F1 scores of both the training and validation sets were used as stopping criteria with a wait time of 20 epochs. So, if the parameters being monitored remain unchanged (i.e. the validation accuracy does not further increase and the validation

loss does not further decrease) after 10 epochs, then the training process will terminate.

On the other hand, learning rate specifies the rate at which the weights are being adjusted. A higher learning rate adjusts the weights more at a time, while a smaller learning rate leads to smaller adjustments. Early into training, a smaller learning rate may not allow the model to achieve its best potential due to convergence to a local minimum instead of a global minimum, while a larger learning rate would allow more rapid optimization. However, as the number of training epochs gets higher, a larger learning rate may negatively affect the optimization by causing non-convergence, while a smaller learning rate can help slowly optimize the model to achieve a global minimum in the loss. Thus, in the study, an exponential decay scheme was used to decrease the learning rate with the increasing number of training epochs.

The models were all trained using ADAM optimizer, and the architectures experimented on are summarized in Table 1, including the pre-trained model architecture, hidden layer architecture, and unfrozen (i.e. trainable) layers, as well as the pre-trained weights.

Web-based application

To allow the conservation community access to the classification tool, a web-based interface was constructed using Streamlit. Streamlit is an open-source platform and Python library that allows machine learning and AI development teams to quickly construct interactive web-based platforms. The models fine-tuned for the classification task were saved as TensorFlow H5 file formats and were reconstructed in the Streamlit environment. The web widgets, including the image upload function, model selection button widgets, and drop-down boxes are built-in functions of the Streamlit library. The interface allows users to upload their microscopic images of mold-contaminated artifacts and classify the strain of the mold

presented. A knowledge base, stored as a Python dictionary, was also constructed to store relevant information about each strain, such as the health risks associated with the strain and recommendations for treatment. Future expansion and refinement of the model can be accelerated with the involvement of the research community. Therefore, a feedback mechanism was implemented to enable researchers and experts to contribute new images to the current image repository, correct labels if the model provides an erroneous classification, and introduce new strains.

Results and discussion

Model performance

The respective performances of the models are summarized in Table 2 corresponding to the models by index number. The parameters recorded to evaluate the model performance were loss, accuracy, precision, recall, and F1 score, for both the training and validation sets. Since early stopping was implemented during the training process, the evaluation parameters are all optimal values that triggered the early termination. More specifically, the maximum values of accuracy, precision, recall, and F1 score, as well as the minimum value of loss are recorded in the table. The epoch at which the training process terminated is recorded in the last column as a reference for convergence speed.

Overall, according to the training performance summarized in Table 2 as well as Fig. 4. Both types of CNN model architectures, VGG16 and ResNet50, demonstrate sufficient convergence between the training and validation sets, indicating the capability of the traditional convolutional neural network for this classification task.

The performance of the developed model was tested using a testing set containing new images of the pre-identified microorganism strains, with the confusion matrix showcased in Fig. 4B. According to the confusion matrix,

Table 1 Modeling Experiment Architectures

Model	Transferred model	Hidden layers	Finetune	Weights
1	VGG16	GAP+FC(5)	All frozen	ImageNet
2	VGG16	Flatten+FC(2048)+FC(1024)+FC(5)	Partial finetune	ImageNet
3	VGG16	GAP+FC(5)	Partial finetune	ImageNet
4	VGG16	GAP+FC(5)	Full finetune	ImageNet
5	VGG16	GAP+FC(5)	Full finetune	Random
6	ResNet50	GAP+FC(5)	All frozen	ImageNet
7	ResNet50	GAP+FC(5)	Partial finetune	ImageNet
8	ResNet50	GAP+FC(5)	Full finetune	ImageNet
9	ResNet50	GAP+FC(5)	Full finetune	Random

The summary of architecture experiments. Transferred Model refers to the main CNN architecture used, Hidden Layers refers to the additional architectures added to the transferred models, Finetune refers to the fine-tuning schemes, and Weights refers to the initialized weights of the pretrained models

Table 2 Model performance summary

Model	Training ¹					Testing ²					Epochs
	Loss	Acc	Precision	Recall	F1	Loss	Acc	Precision	Recall	F1	
1	0.2381	0.9332	0.9572	0.8996	0.9076	0.2868	0.9055	0.9314	0.8702	0.8775	65
2	5.8e-6	1.0	1.0	1.0	0.9864	0.7753	0.8099	0.9160	0.8090	0.7616	36
3	0.0069	0.9990	0.9992	0.9986	0.9876	0.1044	0.9772	0.9784	0.9768	0.9660	53
4	4.7e-6	1.0	1.0	1.0	0.9922	0.1528	0.9788	0.9790	0.9790	0.9715	38
5	0.0007	0.9997	0.9997	0.9997	0.9905	0.0572	0.9902	0.9908	0.9822	0.9785	55
6	0.0983	0.9842	0.9897	0.9754	0.9711	0.1270	0.9663	0.9776	0.9585	0.9607	54
7	7.7e-5	1.0	1.0	1.0	0.9912	0.0016	0.9995	0.9946	0.9946	0.9962	34
8	0.0004	1.0	1.0	1.0	0.9911	0.0057	0.9984	0.9984	0.9984	0.9883	57
9	0.0094	0.9981	0.9981	0.9978	0.9848	0.0159	0.9984	0.9984	0.9984	0.9924	77

The model number corresponds to the models listed in Table 1

¹ The performance in the training phase

² The performance in the testing phase

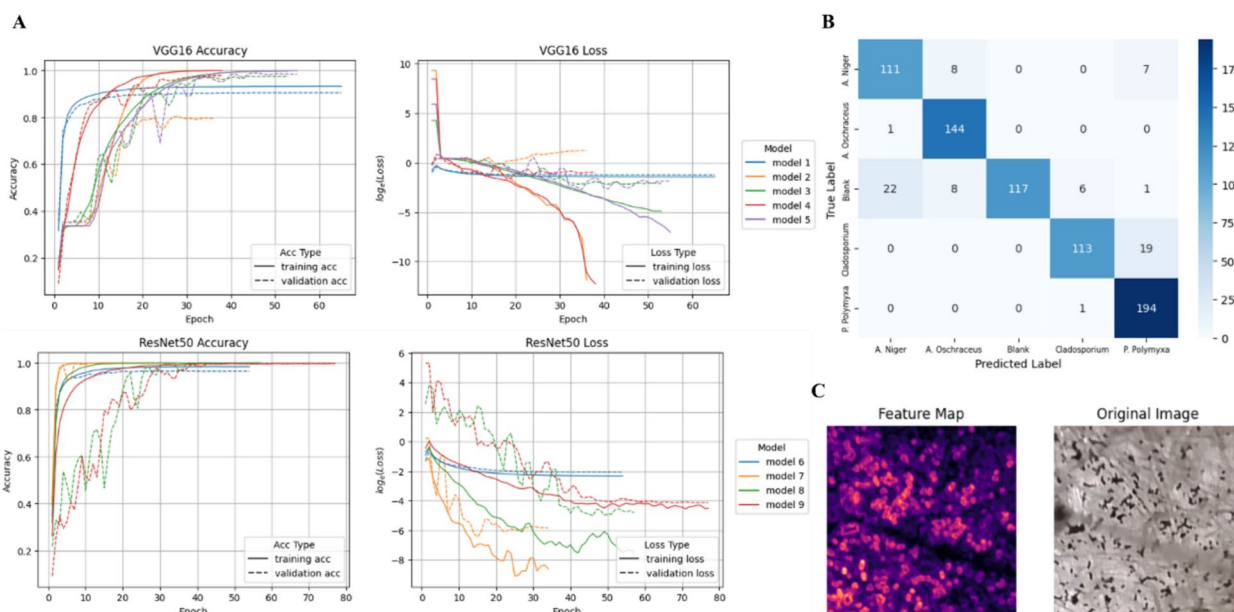


Fig. 4 Model performance summary. **A** Training and validation accuracy and loss for the two major CNN architectures investigated: VGG16 and ResNet50. Training and validation showing convergence tendencies, demonstrating the potential for microscopic image classification. **B** Confusion matrix for the classification accuracy among five categories, including blank control. **C** Feature map of the CNN model

for the total 752 test cases, 679 cases were correctly predicted, achieving a testing accuracy of 90.29%. In detail, *A. Niger* achieved 88.1% accuracy, *A. Oschraceus* achieved 99.3% accuracy, *Cladosporium* achieved 85.6% accuracy, and *P. Polymyxa* achieved the highest prediction accuracy among all four strains of 99.5%. There are 8 false positive predictions of *A. Oschraceus* that should be *A. Niger*, and 1 false positive prediction of *A. Niger* that should be *A. Oschraceus*. These misclassifications could be due to the structural similarity between the two species since they both belong to the same Aspergillus genus.

The misclassification cases of *Cladosporium* concentrate at *P. Polymyxa*, because the two strains share high commonalities in their microscopic features. The blank control has false positive classifications from all of the four strains particularly from *A. Niger*. This could be due to *A. Niger* having the smallest bodies among the four strains, which makes the background paper cellulose structures have more confounding power during the classification.

Further, feature maps of the trained classification models were investigated. Visualizing feature maps in a deep CNN can provide insights into the learned

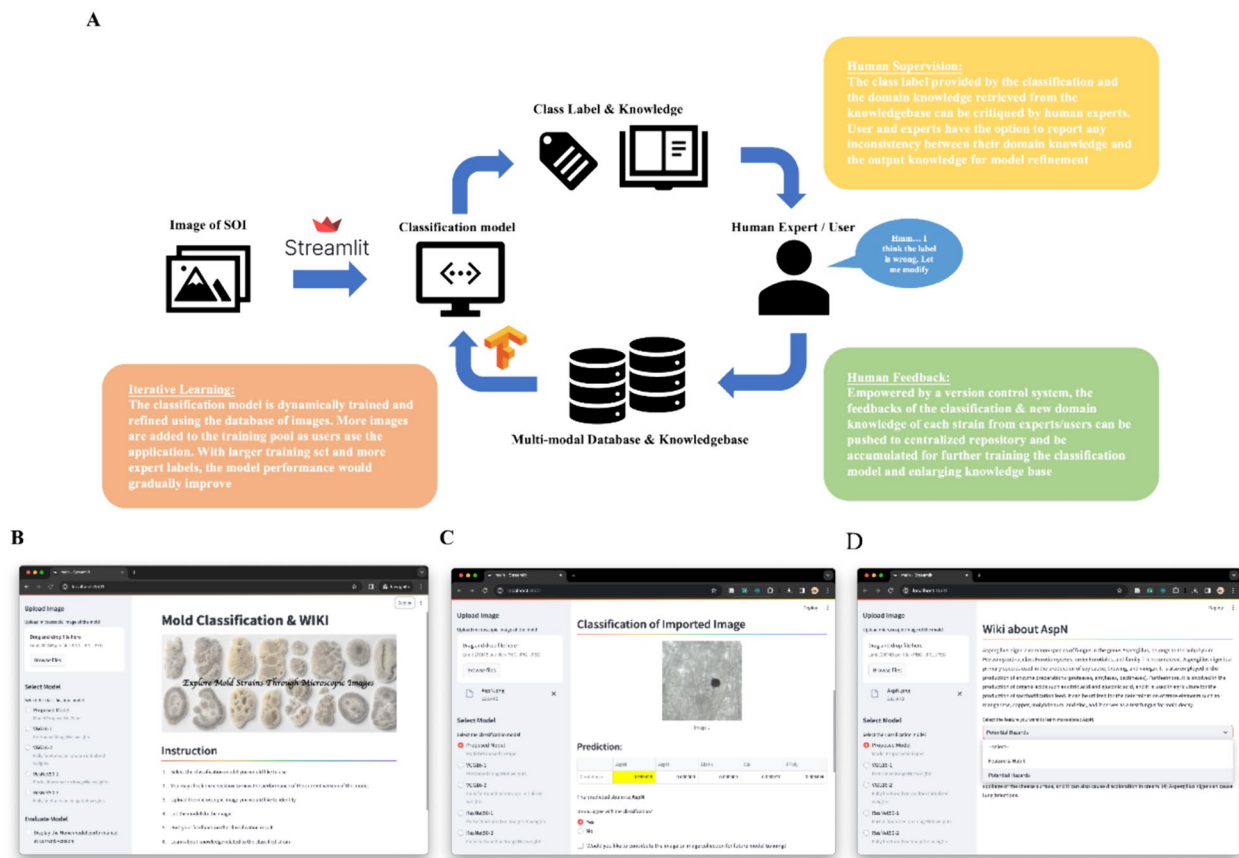


Fig. 5 Mold classification iterative learning pipeline and web-based application. **A** Pipeline for iterative learning of the mold classification model with human feedbacks. **B** Web-based application: main page. **C** Web-based application: classification page. **D** Web-based application: knowledge retrieval and representation wiki page

representations, including patterns, shapes, and textures, to help us understand how the model processes and interprets the input data. Figure 4C shows the feature map vs. the original image for *A. Niger*. The dark dots that scatter the original image are the microorganism bodies and the gray shadows in the background are cellulose fibers. The brighter plasma colors in the feature map are regions that contain the information that the model deems important for the classification task, in other words, the representative features that are learned by the model to identify the respective microorganism. When comparing the feature map with the original image, promising alignments can be seen between the microorganism bodies' contours and the highlighted area on the feature map as well as between the cellulose fibers and the darker shadows in the feature map, indicating the model has successfully learned about the actual morphological features of the

microorganism bodies with contrast to the background cellulose fibrous structures for the classification task.

Web-application & community engagement

The purpose of this segment of the work is to promote resource sharing and establish protocols for the collective advancement of deep learning applications in cultural heritage. As [37] mentioned, the success of deep learning models depends on the power of the dataset that the model was trained on, which includes the number of data and the quality of data. Specifically, a database that incorporates metadata of nucleotide sequences, microbial strains, and potential enzymatic properties as well as their destructive mechanism in the biodeterioration of different materials should be created and curated to assist more efficient identification, removable, and prevention of contaminating fungal strains [9]. At the same time, better resource sharing can be achieved by developing

online platforms. For example, through the collective force of people who share an interest in fungi, a fine-grained classification dataset was developed—Danish Fungi 2020 (DF20) [47]. A mobile application based on CV, named FungiVision, was then established to classify different types of fungi and assist Mycology.

The web-based interface (<https://biodegrade-diagnostics.streamlit.app/>) developed in this study allows three streams of activities that form a closed loop, which we call an “iterative learning scheme” (Fig. 5A). First, when a new image of an unknown strain is uploaded to the application, users can choose from different classification models, which were trained on existing data, to perform the classification task and produce the predicted class label of the unknown strain (Fig. 5B, C). Then, the second stream of functionality is to retrieve the knowledge stored in the knowledge base about the classified strain (Fig. 5D). Currently, the knowledge base that the users can access incorporates information including morphology, cultural and molecular analyses, potential enzymatic properties, and their destructive mechanism to assist practitioners in developing the most suitable treatment techniques. Lastly, the third stream of function allows the user, after receiving classification results from the algorithm, to decide whether to accept the provided label and information. If the human user/expert raises a concern regarding the accuracy of the label provided, they can modify the label or add a new label if the species does not exist in the existing list of strains. The activity, including both the machine output and human input, will be logged in the back-end. Then, users can initiate a pulling request of the administrator to both update the image database and re-train the classification models.

Outlooks & improvements

This project trained a multi-class model that can distinguish between four strains of molds by feeding in images that contain only one strain per image. In real-life situations, such as archives and museums, or even freshly excavated cultural remains, there could be multiple strains of microorganisms contaminating the same surface area of the paper. To adhere to the principle of minimal intervention and apply strain-specific treatments, a multi-label classification model can be further developed, in case multiple different strains of microorganisms are captured in the same microscopic image.

The current project is just the prototype of AI integration in conservation practice. While there are many more strains of molds presented on cultural heritage, the refinement of the model, to let the models become more accurate and more generalizable, requires effort from the research community. The web-based application,

implemented with the human feedback functionality, would serve as a platform to achieve such an endeavor as more and more specialists participate.

Conclusion

Our study showed the competency of CNN models in classifying microscopic images to achieve the end of identifying the microorganism strain presented. A suite of models that are competent in the classification task were generated as deliverables for the conservation community to use and contribute. At the same time, we proposed a sustainable application framework, accessible as a web-based application, to engage the conservation community to collaborate and contribute to the further improvement of the models.

Acknowledgements

Special thanks to the Department of Special Collections of Liaoning University's library for granting access to ancient manuscripts. Thanks to the organizer and participants of the 2023 Art Bio Matters (ABM) conference that provided a platform to share the concept.

Author contributions

Chenshu Liu designed the research study, processed the data used in the study, and developed the software components. Songbin Ben sought funding sources for the project, conducted sample collection, and assisted in the research design. Chongwen Liu helped expand the conservator network and conducted interviews. Xianchao Li, Qingxia Meng, Yilin Hao, and Qian Jiao contributed to the sample collection portion of the project. Pinyi Yang contributed to data processing and assisted with model construction. All authors wrote and revised the main text and contributed to the study conceptualization.

Funding

This project is funded by the Liaoning Province Economic and Social Development Research Project (2023lslybkt-066); Liaoning Provincial Social Science Planning Fund Project (L20BTQ005); Science and Technology Plan Project of Liaoning Provincial Archives Bureau (2021B005).

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Code availability

The code used and/or analyzed during the current study is available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not application.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Bioengineering, University of California, Los Angeles, CA 90024, USA. ²Department of Life Sciences, Liaoning University, Shenyang 110036, Liaoning, China. ³Liaoning Hongshan Culture Heritage Key Laboratory, Liaoning University, Shenyang 110136, Liaoning, China. ⁴UCLA/Getty Interdepartmental Program in the Conservation of Cultural Heritage,

University of California, Los Angeles, CA 90024, USA. ⁵Institute of Data Science, Columbia University, New York, NY 10027, USA.

Received: 30 March 2024 Accepted: 30 April 2024

Published online: 13 May 2024

References

- Sequeira S, Cabrita EJ, Macedo MF. Antifungals on paper conservation: an overview. *Int Biodeterior Biodegrad*. 2012;74:67–86. <https://doi.org/10.1016/j.ibiod.2012.07.011>.
- Daniels V. The chemistry of paper conservation. *Chem Soc Rev*. 1996;25(3):179–86. <https://doi.org/10.1039/C59962500179>.
- Trovão J, Portugal A. Current knowledge on the fungal degradation abilities profiled through biodeteriorative plate essays. *Appl Sci*. 2021;11(9):4196. <https://doi.org/10.3390/app11094196>.
- Campbell AW. Molds and mycotoxins: a brief review. *Altern Ther Health Med*. 2016;22(4):8–11.
- Anton R, Moularat S, Robine E. A new approach to detect early or hidden fungal development in indoor environments. *Chemosphere*. 2016;143:41–9. <https://doi.org/10.1016/j.chemosphere.2015.06.072>.
- Pinheiro AC, Sequeira SO, Macedo MF. Fungi in archives, libraries, and museums: a review on paper conservation and human health. *Crit Rev Microbiol*. 2019;45(5–6):686–700. <https://doi.org/10.1080/1040841X.2019.1690420>.
- Montanari M, Melloni V, Pinzari F, Innocenti G. Fungal biodeterioration of historical library materials stored in Compactus movable shelves. *Int Biodeterior Biodegrad*. 2012;75:83–8. <https://doi.org/10.1016/j.ibiod.2012.03.011>.
- Tahir MW, Zaidi NA, Rao AA, Blank R, Vellekoop MJ, Lang W. A fungus spores dataset and a convolutional neural network based approach for fungus detection. *IEEE Trans Nanobiosci*. 2018;17(3):281–90. <https://doi.org/10.1109/TNB.2018.2839585>.
- Sterflinger K, Little B, Pinar G, Pinzari F, Rios A, Gu JD. Future directions and challenges in biodeterioration research on historic materials and cultural properties. *Int Biodeterior Biodegrad*. 2018;129:10–2. <https://doi.org/10.1016/j.ibiod.2017.12.007>.
- Florian ML. Review of fungal facts: solving fungal problems in heritage collections. *J Am Inst Conserv*. 2004;43(1):114–6. <https://doi.org/10.2307/3179856>.
- Meade E, Slattery MA, Garvey M. Biocidal resistance in clinically relevant microbial species: a major public health risk. *Pathogens*. 2021;10(5):598. <https://doi.org/10.3390/pathogens10050598>.
- Sully D. Conservation theory and practice: materials, values, and people in heritage conservation. In: Macdonald S, Leahy HR, editors. *The International handbooks of museum studies*. Hoboken: Wiley; 2015. <https://doi.org/10.1002/9781118829059.wbihsms988>.
- Vettraino AM, Zikeli F, Humar M, Biscontri M, Bergamasco S, Romagnoli M. Essential oils from *Thymus* spp as natural biocide against common brown- and white -rot fungi in degradation of wood products: antifungal activity evaluation by in vitro and FTIR analysis. *Eur J Wood Wood Prod*. 2023;81:747–63. <https://doi.org/10.1007/s00107-022-01914-3>.
- Rahmani TPD, Ismail I, Aziz IR. Biodeterioration and biodegradation of cultural & religious heritage made of paper as a wood derivative. *J Islam Sci*. 2022;9(1):52–7. <https://doi.org/10.24252/jis.v9i1.30285>.
- Florian ML, Koestler RJ, Nicholson K, Parker TA, Stanley T, Szczepanowska H, Wagner S. Chapter 12 Mold/fungi. S. Bertalan, editors. *Paper Conservation Catalog*. 9th edition. 1994. https://cool.culturalheritage.org/coolaic/sg/bpg/pcc/1994_frontmatter.pdf. Accessed 8 July 2023.
- Nguyen LD, Lin D, Lin Z, Cao J. Deep CNNs for microscopic image classification by exploiting transfer learning and feature concatenation. *IEEE International Symposium on Circuits and Systems (ISCAS)*. 2018; 1–5. <https://doi.org/10.1109/ISCAS.2018.8351550>.
- Liu L, Li Y, Li S, Hu N, He Y, Pong R, Lin D, Lu L, Law M. Comparison of next-generation sequencing systems. *J Biomed Biotechnol*. 2012;2012:251364. <https://doi.org/10.1155/2012/251364>.
- Quail MA, Smith M, Coupland P, Otto TD, Harris SR, Connor TR, Bertoni A, Swerdlow HP, Gu Y. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics*. 2012;13:341. <https://doi.org/10.1186/1471-2164-13-341>.
- Voulodimos A, Doulamis N, Doulamis A, Protopapadakis E. Deep learning for computer vision: a brief review. *Comput Intell Neurosci*. 2018;2018:7068349. <https://doi.org/10.1155/2018/7068349>.
- Esteve A, Chou K, Yeung S, Naik N, Madani A, Mottaghi A, Liu Y, Topol E, Dean J, Socher R. Deep learning-enabled medical computer vision. *Npj Digit Med*. 2021;4:5. <https://doi.org/10.1038/s41746-020-00376-2>.
- Zieliński B, Sroka-Oleksiak A, Rymarczyk D, Piekarczyk A, Brzychczy-Włoch M. Deep learning approach to describe and classify fungi microscopic images. *PLoS ONE*. 2020;15(6):0234806. <https://doi.org/10.1371/journal.pone.0234806>.
- Hatir ME, Barstuğan M, İnce İ. Deep learning-based weathering type recognition in historical stone monuments. *J Cult Herit*. 2020;45:193–203. <https://doi.org/10.1016/j.culher.2020.04.008>.
- Zhang Y, Li J, Tang F, Zhang H, Cui Z, Zhou H. An automatic detector for fungal spores in microscopic images based on deep learning. *Appl Eng Agric*. 2021;37(1):85–94. <https://doi.org/10.13031/aea.13818>.
- Ngo CC, Nguyen QH, Nguyen TH, Quach NT, Dudhagara P, Vu THN, Le TTX, Le TTH, Do TTH, Nguyen VD, Nguyen NT, Phi QT. Identification of fungal community associated with deterioration of optical observation instruments of museums in Northern Vietnam. *Appl Sci*. 2021;11(12):5351. <https://doi.org/10.3390/app11125351>.
- Romero SM, Giudicessi SL, Vitale RG. Is the fungus *Aspergillus* a threat to cultural heritage? *J Cult Herit*. 2021;51:107–24. <https://doi.org/10.1016/j.culher.2021.08.002>.
- Kosel J, Ropret P. Overview of fungal isolates on heritage collections of photographic materials and their biological potency. *J Cult Herit*. 2021;48:277–91. <https://doi.org/10.1016/j.culher.2021.01.004>.
- Karbowska-Berent J, Górniak B, Czajkowska-Wagner L, Rafalska K, Jarmińko J, Kozięlec T. The initial disinfection of paper-based historic items – observations on some simple suggested methods. *Int Biodeterior Biodegrad*. 2018;131:60–6. <https://doi.org/10.1016/j.ibiod.2017.03.001>.
- Boniek D, Bonadio L, Damaceno QS, Santos AFB, Resende Stoianoff MA. Occurrence of *aspergillus niger* strains on a polychrome cotton painting and their elimination by anoxic treatment. *Can J Microbiol*. 2020;66(10):586–92. <https://doi.org/10.1139/cjm-2020-0173>.
- Carvalho HP, Mesquita N, Trovão J, Silva JP, Rosa B, Martins R, Bandeira AML, Portugal A. Diversity of fungal species in ancient parchments collections of the archive of the University of Coimbra. *Int Biodeterior Biodegrad*. 2016;108:57–66. <https://doi.org/10.1016/j.ibiod.2015.12.001>.
- Pavlović J, Farkas Z, Kraková L, Pangallo D. Color stains on paper: fungal pigments, synthetic dyes and their hypothetical removal by enzymatic approaches. *Appl Sci*. 2022;12(19):9991. <https://doi.org/10.3390/app12199991>.
- Rojas TI, Aira MJ, Batista A, Cruz IL, González S. Fungal biodeterioration in historic buildings of Havana (Cuba). *Grana*. 2012;51(1):44–51. <https://doi.org/10.1080/00173134.2011.643920>.
- Michaelsen A, Piñar G, Pinzari F. Molecular and microscopical investigation of the microflora inhabiting a deteriorated Italian manuscript dated from the Thirteenth Century. *Microb Ecol*. 2010;60(1):69–80. <https://doi.org/10.1007/s00248-010-9667-9>.
- Lech T. Evaluation of a parchment document, the 13th century incorporation charter for the City of Krakow, Poland, for microbial hazards. *Appl Environ Microbiol*. 2016;82(9):2620–31. <https://doi.org/10.1128/AEM.03851-15>.
- Principi P, Villa F, Sorlini C, Cappitelli F. Molecular studies of microbial community structure on stained pages of Leonardo da Vinci's Atlantic Codex. *Microb Ecol*. 2011;61(1):214–22. <https://doi.org/10.1007/s00248-010-9741-3>.
- Chan HP, Hadjiiski LM, Samala RK. Computer-aided diagnosis in the era of deep learning. *Med Phys*. 2020;47(5):218–27. <https://doi.org/10.1002/mp.13764>.
- Kim HE, Cosa-Linan A, Santhanam N, Jannesari M, Maros ME, Ganslandt T. Transfer learning for medical image classification: a literature review. *BMC Med Imaging*. 2022;22:69. <https://doi.org/10.1186/s12880-022-00793-7>.
- Fujita H. AI-based computer-aided diagnosis (AI-CAD): the latest review to read first. *Radiol Phys Technol*. 2020;13(1):6–19. <https://doi.org/10.1007/s12194-019-00552-4>.

38. Deng J, Dong W, Socher R, Li LJ, Li K, Li FF. Imagenet: a large scale hierarchical image database. *IEEE Conference on Computer Vision and Pattern Recognition*. 2009; 248–255. <https://doi.org/10.1109/CVPR.2009.5206848>
39. Liu Z, Jin L, Chen J, Fang Q, Ablameyko S, Yin Z, Xu Y. A survey on applications of deep learning in microscopy image analysis. *Comput Biol Med*. 2021;134:104523. <https://doi.org/10.1016/j.combiomed.2021.104523>.
40. Alzubaidi L, Fadhel MA, Al-Shamma O, Zhang J, Duan Y. Deep learning models for classification of red blood cells in microscopy images to aid in sickle cell anemia diagnosis. *Electronics*. 2020;9(3):427. <https://doi.org/10.3390/electronics9030427>.
41. Chai J, Zeng H, Li A, Ngai EWT. Deep learning in computer vision: a critical review of emerging techniques and application scenarios. *Mach Learn Appl*. 2021;6:100134. <https://doi.org/10.1016/j.mlwa.2021.100134>.
42. Salehi AW, Khan S, Gupta G, Alabdullah BI, Almjally A, Alsolai H, Siddiqui T, Mellit A. A study of CNN and transfer learning in medical imaging: advantages, challenges, future scope. *Sustainability*. 2023;15(7):5930. <https://doi.org/10.3390/su15075930>.
43. Cebollada S, Payá L, Flores M, Peidró A, Reinoso O. A state-of-the-art review on mobile robotics tasks using artificial intelligence and visual data. *Expert Syst Appl*. 2021;167:114195. <https://doi.org/10.1016/j.eswa.2020.114195>.
44. Mascarenhas S, Agarwal M. A comparison between VGG16, VGG19 and ResNet50 architecture frameworks for Image Classification. *International conference on disruptive technologies for multi-disciplinary research and applications (CENTCON)*. 2021; 1: 96-99. <https://doi.org/10.1109/CENTCON52345.2021.9687944>
45. Lin M, Chen Q, Yan S. Network in network. *arXiv preprint arXiv:1312.4400*. 2013. <https://doi.org/10.48550/arXiv.1312.4400>
46. Zhou B, Khosla A, Lapedriza A, Oliva A, Torralba A. Learning Deep Features for Discriminative Localization. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016; 2921–2929. <https://doi.org/10.1109/CVPR.2016.319>
47. Picek L, Šulc M, Matas J, Heilmann-Clausen J, Jeppesen TS, Lind E. Automatic fungi recognition: deep learning meets mycology. *Sensors*. 2022;22(2):633. <https://doi.org/10.3390/s22020633>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.