Kemal Hacıefendioğlu[1], Hasan Basri Başağa[1], Murat Emre Kartal[2], Mehmet Ceyhun Bulut[1]

Automatic Damage Detection on Traditional Wooden Structures with Deep Learning-Based Image Classification Method

Automatsko otkrivanje oštećenja na tradicionalnim drvenim konstrukcijama metodom klasifikacije slika utemeljenom na dubokom učenju

ORIGINAL SCIENTIFIC PAPER
Izvorni znanstveni rad
Received – prispjelo: 25. 2. 2021.
Accepted – prihvaćeno: 13. 12. 2021.
UDK: 624.011.1
https://doi.org/10.5552/drvind.2022.2108

ABSTRACT • *Wood has a long history of being used as a valuable resource when it comes to building materials. Due to various external factors, in particular the weather, wood is liable to progressive damage over time, which negatively impacts the endurance of wooden structures. Damage assessment is key in understanding, as well as in effectively mitigating, problems that wooden structures are likely to face. The use of a classification system, via deep learning, can potentially reduce the probability of damage in engineering projects reliant on wood. The present study employed a transfer learning technique, to achieve greater accuracy, and instead of training a model from scratch, to determine the likelihood of risks to wooden structures prior to project commencement. Pretrained MobileNet_V2, Inception_V3, and ResNet_V2_50 models were used to customize and initialize weights. A separate set of images, not shown to the trained model, was used to examine the robustness of the models. The three models were compared in their abilities to assess the possibilities and types of damage. Results revealed that all three models achieve performance rates of similar reliability. However, when considering the loss ratios in regard to efficiency, it became apparent that the multi-layered MobileNet_V2 classifier stood out as the most effective of the pre-trained deep convolutional neural network (CNN) models.*

KEYWORDS: *deep learning method; convolutional neural networks; MobileNet_V2; Inception_V3; ResNet_V2_50; wooden structures*

SAŽETAK • *Drvo kao vrijedan građevni materijal ima dugu povijest uporabe u graditeljstvu. No zbog brojnih vanjskih čimbenika, posebice vremenskih utjecaja, drvo tijekom vremena postaje podložno progresivnom propadanju, što negativno utječe na izdržljivost drvenih konstrukcija. Procjena šteta na drvu ključna je za razumijevanje*

[1] Authors are researchers at Karadeniz Technical University, Department of Civil Engineering, Trabzon, Turkey.
[2] Author is researcher at İzmir Democracy University, Department of Civil Engineering, İzmir, Turkey.

problema koji će vjerojatno nastati na drvenim konstrukcijama, kao i za njihovo učinkovito ublažavanje. Primjena sustava klasifikacije uz pomoć dubokog učenja može potencijalno smanjiti vjerojatnost oštećenja u inženjerskim projektima koji se oslanjaju na drvo. U ovom je istraživanju primijenjena tehnika transfernog učenja kako bi se postigla veća točnost modela umjesto da se model za utvrđivanje vjerojatnosti rizika za drvene konstrukcije radi prije početka projekta. Za prilagodbu i inicijalizaciju težina primijenjeni su unaprijed osposobljeni modeli MobileNet_V2, Inception_V3 i ResNet_V2_50. Za ispitivanje robusnosti modela upotrijebljen je zaseban skup slika koji nije prikazan u osposobljenome modelu. Spomenuta tri modela uspoređena su s obzirom na njihove mogućnosti procjene vjerojatnosti i vrste oštećenja drvenih konstrukcija. Rezultati su otkrili da sva tri modela imaju sličnu pouzdanost. Međutim, kada se uzmu u obzir omjeri gubitaka u odnosu prema učinkovitosti, postalo je očito da se višeslojni MobileNet_V2 klasifikator istaknuo kao najučinkovitiji od unaprijed pripremljenih modela dubokih konvolucijskih neuronskih mreža (CNN).

KLJUČNE RIJEČI: *metoda dubokog učenja; konvolucijska neuronska mreža; MobileNet_V2; Inception_V3; ResNet_V2_50; drvene konstrukcije*

1 INTRODUCTION
1. UVOD

Wooden structures have a long and varied history in the realms of engineering. Indeed, throughout the world, and specifically in Turkey, a number of very old wooden structures have managed to survive. It is important for cultural posterity that these structures be protected, however, this can only be achieved with adequate resources, including the ability to predict potential risks in regard to damage, and how best to carry out repairs. Intensive studies regarding longevity estimation and evaluation methodologies have been carried out in recent years with a view to ensuring sustainability of older wooden structures, while also negotiating the environmental and economic impact of conservation. When evaluating the long-term performance of buildings and building materials, a key aim is to reduce the ecological and economic effects brought about by maintenance, repairs, and renewal, and thus use resources effectively (Jones *et al*., 2007). It is essential to know the environmental impact that building materials and elements can cause throughout the entire service life. An examination of the available research evidently showed that potential damage to building elements and materials are best examined through the preparation of a condition evaluation protocol and subsequent processing on a mapping system. This system, known as a damage map, is the sum of studies that show deterioration in various building elements via dimages, which can then be used to explain the causes of deterioration and the specific conditions that result in certain types of damage determining their degree, and provide environmentally and economically sustainable repair suggestions. Damage that may occur over time due to environmental factors in wooden building elements and materials are usually categorized as surface changes, color changes, fragmentation, cracks, and mechanical and biological degradations. In this study, the damage that occurs in the building elements of wooden structures are examined within three categories, namely wet rot damage, dry rot damage, and insect damage.

Dry rot damage is a wood-destroying fungus that causes timber rot by feeding wood cells and reducing their strength. This type of damage is a term used to describe any kind of decay that causes deterioration resulting in blackened, weakened, and cracked wood in buildings and ships caused by the fungus (Robinson, 2005). Dry rot damage is the most severe form of fungal rot. It attacks timber in buildings, and essentially digests parts of the wood, weakenng it severely over time. The scientific name of the fungus that creates dry rot is *Serpula Lacrymans*. Since wood can produce moisture through digestion, it can spread without any source of moisture. When dry rot damage spreads, it can seriously damage the structural integrity of the building. If dry rot damage is not immediately detected and treated, all affected timber may need to be removed and replaced. Spores of the fungus that create dry rot are found in the atmosphere; however, the right conditions for germination are needed for the fungus to take hold. These conditions include moist timber with a moisture content of about 20 %, which is freely accessible to the air (Jones *et al*., 2007). Wet rot damage is caused by a fungus that affects timber in very humid environments. High humidity creates rot by attracting spores that produce excess moisture. Compared to dry rot damage, wet rot damage is less destructive because wet rot damage remains limited to wet areas. However, it must still be treated effectively so as not to erode the structural integrity of wooden buildings (Jones *et al*., 2007). Wet rot damage can cause significant structural damage if left to grow uncontrollably because it will weaken timber over time. Wet rot is a general term for several fungi types, the most common being *Coniophora Puteana*, also known as cellar fungus. Wet rot damage is formed when the lumber is exposed to excessive moisture over longer periods of time. The wet timber will continue to soften as it absorbs more and more moisture thus leading to significant damage (Jones *et al*., 2007). This type of damage is often found in cellars, roofs, and window casings. The danger of wet rot damage when left untreated is the potential significant erosion of a wooden structure strength.

Insect damage can also be a common problem that affects wooden structures. Damage is typically caused by termites, carpenter ants, and powder insects. Although termites usually cause the most significant damage, carpenter ants and powder insects must also be accounted for. Like termites, carpenter ants carve channels through wood, however, unlike their counterparts, the grooves of carpenter ants have a smooth and clean appearance (Robinson, 2005). Although damage inflicted by carpenter ants is typically not as severe as the damage caused by termites, carpenter anthills can cause significant damage if left untreated over the years. As the number of individuals in a colony increases, they will continue to break down the wood, causing further damage (Jones *et al*., 2007). The larvae of powdered insects feed on cellulose in wood and cause significant structural damage over a period of one to five years. As the insects feed, a fine powder in the cavities under the wooden surface is produced. They typically inhabit soft woods and are therefore less likely to be the cause of substantial structural problems (Robinson, 2005), however, their effects should not be discounted.

With the development of technology in civil engineering, many preliminary examinations such as damage inspection, detection of defects, and determination of material strengths are now carried out with computers to evaluate structural performance against external influences. In wooden structures, it is essential to detect any damage or defects of wooden materials and, in light of this, a number of studies have been carried out over the last few decades. One of the first of these kinds of studies is the detection and classification of wood defects using artificial intelligence performed by Cavalin *et al*. (2006). They attempted to detect wood defects using features determined from grayscale images. They defined features according to smoothness, coarseness, and regularity. In experimental studies, neural networks and support vector machines with two different learning paradigms were taken into account, as well as a feature selection algorithm based on multi-objective genetic algorithms. The experimental results indicate that after feature selection the grayscale image-based feature set accomplishes superior wood defect detection performance than color image-based features. In a study carried out by Jabo (2011), image analysis and classification methods were proposed. These methods cover the sorting of defects, including "wood decay", "blue stain", and "shake". A supervised classifier was first trained with Adaboost, and used to extract colors of stain-type defects and methods such as an integral image. Sioma (2015) presented the analysis and use of 3D images to detect and locate defects on wood surfaces automatically. In the study, defects occurring on wood surfaces were explained in detail, and examples were presented of the most common defects recorded in wood industry production lines. A method was applied to create a 3D image of the surface using the laser triangulation method (LTM). For selected defects, measurement algorithms that can detect a parametric description of the defect and its location on the surface were presented. Wenshu *et al*. (2015) aimed to identify the location and size of defects on wood quickly and accurately by using a computer and artificial neural network (ANN) technology. In their work, the location and size of the defects in the wood were determined by taking images with a CCD (charged coupling devices) camera, and processing said images using the MATLAB program. Additionally, ANN was created to recognize defects with results showing that the wood defect identification rate is ascertainable up to 86.67 %. Urbonas *et al*. (2019) proposed an automatic visual inspection system for the location and classification of defects on wooden surfaces. A faster region-based CNN (Faster R-CNN) method was used to identify defects in wood veneer surfaces. Pre-trained AlexNet, VGG16, BNInception and ResNet152 neural network models were used for transfer learning to improve the results. In their study, Li *et al*. (2019) proposed a new methodology, including the application of machine learning algorithms, to cope with damage accumulation effects in wood. The proposed algorithm takes into account a multi-objective optimization process with a combination of goodness of fit and complexity.

Researchers have used long-term experimental data of typical wood species to develop damage accumulation models based on machine learning. He *et al*. (2019) and (2 proposed a diverse, fully convolutional neural network (Mix-FCN) to locate wood defects and automatically classify defect types in wood surface images. The images were first collected with a data acquisition device developed in the laboratory. They then utilized TensorFlow and Python language (Van Rossum and Drake Jr, 1995) to create a VGG16 model. They used two types of data sets (dataset 1 and dataset 2) to maximize the data, which was collected on a limited basis, to ensure that the Mix-FCN merges quickly during training. The proposed models were trained, validated, and tested by dataset 2. In order to classify lumber images, Hu *et al*. (2019) explored varieties of deep learning strategies based on ResNet18. Their study four datasets were manually marked as lumber defects, wood textures, and lumbers by experts. In the study, they used transfer learning in CNN with a classifier layer that provides training with a small amount of training data. The accuracy rate was clearly upgraded by expanding unbalanced samples. He *et al*. (2020) suggested a learning method to perceive wood defects as well as automatically categorize imperfections in wood images obtained using a laser scanner via a deep convolutional neural network (DCNN). They implemented TensorFlow to train the network composed of an input layer, four convolutional layers, four maximum pooling layers, three fully connected layers, a softmax layer, and an output layer. Results indicated that the DCNN model could identify wood defects more accurately and effectively than traditional methods.

Turkey, particularly the Anatolian region, has relied on the use of local materials, including wood, stone, and adobe, for much of its history to this day. Wood is used in a multitude of ways as follows: as a carrier element, for exterior cladding, joinery, flooring, and roofing material in various buildings from past to present. Although foundation and basement walls are typically made of stone along the coastline of Turkey, the upper floors are typically constructed of wooden masonry or have a wooden frame system. Serenders are the most common structures in which wooden masonry and a carcass system are jointly used. Serenders are warehouses where various food items are placed. Unlike dwellings, although the lower floor is carcass with buttress poles, the upper floor is wooden masonry. Regardless of the materials used for a building exterior walls, the interior walls, floors, and ceilings are made of wood (Özgel Felek, 2020).

Object recognition is conducted by looking at the attributes associated with an object and identifying criteria that can determine appropriate labeling. Recognition comes from experience, education, and training. Deep learning methods allow programs to quickly recognize and classify an object while taking into account features that the human eye cannot distinguish. The deep learning-based classification methods used in the present study are designed for the visual identification of types of damage. Typically, dry damage and insect damage are difficult to distinguish with the human eye as they often appear similar in texture. While the human eye can certainly detect specific classifications of damage, deep learning allows for classification involving more complex features, and the ability to do so with greater speed and practicality.

2 AIM OF STUDY
2. CILJ ISTRAŽIVANJA

The present study sought to make use of CNN based on the deep learning method, to better automate the assessment of traditional wooden structures in Turkey. The primary focus was the automatic detection and localization of actual damage caused by environmental influences on structural wooden elements via the use of images processed through deep learning techniques. It is important to note that, as this was the first attempt at such an assessment, there were a few limitations. Firstly, more than one type of damage was not taken into account at the same time, which means that the images used by the model belong only to one category. Secondly, only images with visible damage were taken into account. Thirdly, excessive or low illumination affecting the brightness of images as well as the orientation of the images needed to be accounted for.

The aim of the present study is to automatically detect partial damage in wooden structures, which are widely used in Turkey, by using image classification methods based on deep learning. For this purpose, a partially damaged selender structure was used as the example structure.

3 DEEP LEARNING METHOD
3. METODA DUBOKOG UČENJA

Deep learning is a machine learning technique that identifies features and tasks directly from data. Data typically consists of sounds, images, or text. Deep learning is a systematic subset of artificial intelligence and machine learning, and is the most popular artificial intelligence application approach. Studies can be referred to as data interpretation when it comes to the generalization of deep learning (LeCun *et al*., 2015). Deep learning algorithms take all the pixels on the pattern as input data when learning through a tagged pattern. The input data varies depending on whether the pattern is gray or colored. When a color image is given as input to the model, whereby red, green, and blue frequencies are used, the current pixel is used three times as input data. In a grayscale image, the input data is the number of pixels. Primarily, convolution layers, ReLu (Rectified Linear Unit) layers, and pooling layers are applied in order to draw the feature map on the image. After the feature map is drawn, fully linked layers and unique classification layers of Softmax come into play. At this stage, each pixel has probability values for each class. These values are determined by Softmax, which identifies the class of the pixels (Hassaballah and Awad, 2020).

3.1 Convolutional neural networks
3.1. Konvolucijska neuronska mreža

Artificial neural networks are models created based on the functioning of the human brain. The system aims to realize the learning process through the interpretation of these teachings with the system then making decisions autonomously due to this interpretation. This structure consists of three layers. Observations are transferred to the system in the input layer. The number of nodes in this layer is equal to the number of features that will best represent the observation. Values can be copied and sent to more than one layer in the input layer. The hidden layer applies transformations to the values from the input layer by multiplying them with specific coefficients. There can be more than one node in this layer. By using certain threshold conditions at these nodes, a value as much as the output number can be obtained after the hidden layer is the output layer. The system makes predictions according to the value obtained.

CNN is a class of deep neural networks used in image recognition problems (Salman *et al*., 2018)such as Dynamic Movement Primitives (DMPs. According to CNN working principle, the pictures given as input must be recognized by computers and converted into an operable format. For this reason, images are first converted to matrix format.

CNNs are structures designed to take images as inputs and are thus used effectively in computer vision. CNN consists of one or more complex layers and fully bonded layers, such as a standard multilayer neural network (LeCun *et al*., 2015). CNN architecture consists of convolution, pooling, flattened and fully connected layers dropout and classification as shown in Figure 1. If the CNN architecture is considered as two separate sections, and feature extraction for the unit where convolution and pooling layers are used, the unit, where flattening and a fully bonded layer is used, is regarded as the classification. The section after the flattening layer has the same structure as standard neural networks.

Convolution layers are the central part in which the object recognition problem calculations are the most intense. Raw data is given as input to this unit. In this layer, many convolution layers are interconnected to perform a series of convolution operations. Each convolution layer represents a series of feature maps with organized neuron cells. In other words, the attributes of the image given in input are extracted in these layers. The parameters of the layers are a set of learnable filter coefficients. As the convolution layers progress from input to output, features, from simple to complex, are extracted.

The convolutional layers create feature maps with linear filters followed by linear functions (rectifier, sigmoid, tanh, etc.). The feature map can be calculated using Eq. 1:

$$f_{i,j,k} = max(w_k^T x_{i,j} + b_k, 0)$$

Where (i, j) defines the spatial location on feature maps, k is the feature map index, $x_{i,j}$ denotes the input patch centered at (i, j) and w_k^T, b_k are the filter coefficient vector and bias of the k_{th} feature map, respectively. In CNN architecture, a pooling layer is commonly added between the convolution layers periodically. This layer function gradually decreases the notation spatial dimension to reduce the network's number of parameters and calculations, which controls excessive compliance. This is known as a sub-sampling. The pooling layer makes the sample by lowering the feature map height and width although the depth remains the same. Spatial pooling has types such as maximum pooling, average pooling, and total pooling. The most used technique is maximum pooling. Maximum pooling takes the largest item in the corrected feature map. The size of the feature map is halved after pooling.

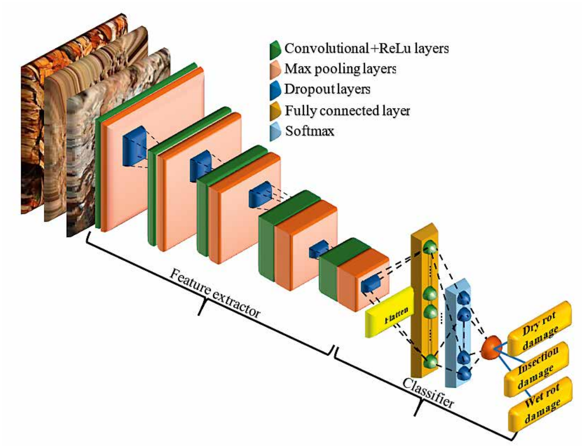

Figure 1 Typical convolutional neural network
Slika 1. Tipična konvolucijska neuronska mreža

Neurons in a fully-connected layer, as in regular neural networks, have full connections with all activations in the previous layer. It is the last and most crucial layer of CNN. It is identified as fully connected because every neuron in the previous layer is connected to each neuron in the next layer. The fully-connected layer aims to use high-level features to classify the input image into different classes. As seen in Figure 4, there are four networks in the fully connected layer. Each of these neurons represents a class (e.g. house, boat, tree, cat, etc.). By looking at the convolutional network attributes, it detects which class the input data is closer to and ultimately returns a class (Schmidhuber, 2015).

3.2 Transfer learning
3.2. Transferno učenje

Transfer learning uses deep learning methods to hide the information obtained while solving a problem and then uses that information when faced with another problem. When it comes to the transfer of learning, models that result in higher success rates and via faster learning with less training data, are typically obtained by using previous knowledge. It is typical for a separate 'learning from scratch' to be performed for each task in deep learning. Since it is possible and advantageous to use information learned from some tasks in other tasks, the source task information is used to solve the target task. Features, weights, etc., are obtained from previously trained models with transfer learning used for a new task. For this method to work, the information transferred needs to be general information.

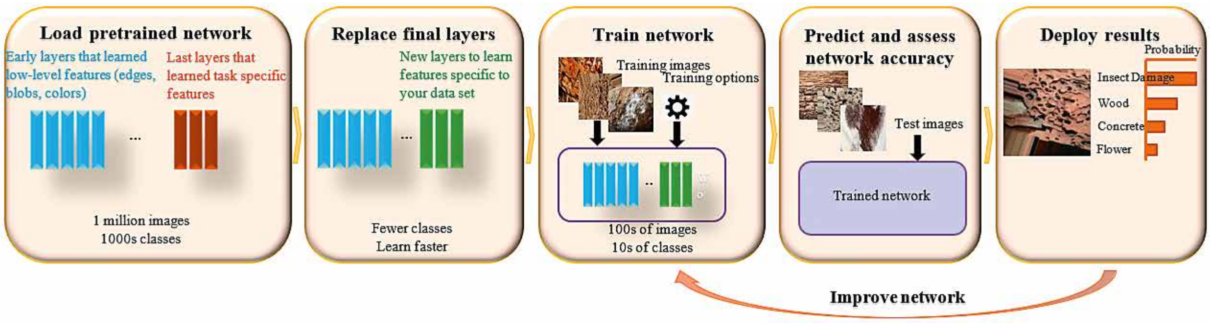

Figure 2 Reuse of pre-trained network
Slika 2. Ponovna uporaba unaprijed osposobljene mreže

Appropriate information for both the source and target tasks is transmitted instead of specific to the source task. It is common practice to purchase and use ready-trained models such as VGG-16, VGG-19, ResNet, Inception, and Xception with Keras library. Pan and Yang (2010) provide a comprehensive review of transfer learning. The reuse of pre-trained network is shown in Figure 2. In this study, Resnet-50, VGG-19, and Inception-V3 models were utilized to localize landslide regions, where wooden structures are particularly liable to damage.

3.2.1 Trained convolutional neural networks
3.2.1. Osposobljena konvolucijska neuronska mreža

The present study used three pre-trained convolutional neural network models -MobileNet_V2, Inception_V3, and ResNet_V2_50, to perform classification on a new collection of images.

3.2.1.1 MobileNet_V2 model
3.2.1.1. Model MobileNet_V2

MobileNet_V2 consists of two types of blocks as shown in Figure 3. The first is a residual block with a stride of 1 and the second is a block with a stride of 2 for downsizing. Three layers are defined for both block types. The first layer is 1×1 convolution with ReLU6. The second layer is representative of the depthwise convolution. The third layer is 1×1 convolution but has non-linearity. In the case of ReLU, the deep networks have a linear classifier power only in the non-zero volume portion of the output domain. There is an expansion factor *t*, and *t*=6 for all primary experiments. If the input had 64 channels, for example, the internal output would have 64×*t*=64×6=384 channels (Sandler *et al*., 2018). For this model, the input images sizes were fixed at a height x width = 224 × 224 pixels.

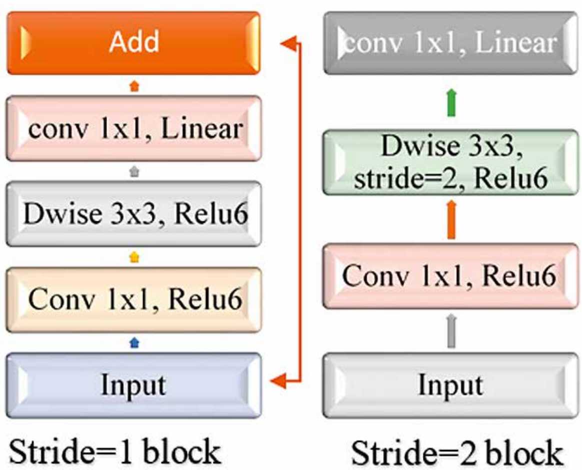

Figure 3 MobileNet_V2 convolutional blocks
Slika 3. Konvolucijski blokovi modela MobileNet_V2

3.2.1.2 Inception_V3 model
3.2.1.2. Model Inception_V3

The inception network model consists of various modules. Different sizes of convolution and maximum partnering are carried out in each of these modules. GoogLeNet is also known as Inception_V1 in the literature (Szegedy *et al*., 2015). Versions later called Inception_V2 (Szegedy *et al*., 2016), Inception_V3 (Szegedy *et al*., 2016), and Inception_V4 (Szegedy *et al*., 2017) have also been developed. Inception_V3 consists of two parts: feature extraction and classification. The feature extraction section involves the convolutional neural network. Inception_V3 is a familiar architecture, and the input of the network should be an image of 299×299 pixels. Moreover, the classification section is wholly connected and contains Softmax layers. All layers specified in Inception_V3 are depicted in Figure 4.

3.2.1.3 ResNet_V2_50 model
3.2.1.3. Model ResNet_V2_50

ResNet_V2_50 (He *et al*., 2015) is a convolutional neural network that is 50 layers deep with a pre-trained version of the network trained on more than one million images from the ImageNet database. The pre-trained network can classify images into 1000 object categories (e.g. keyboard, computer mouse, pencil, animal, etc.). As a result, the network has learned rich feature representations for a wide range of images. The network has an image input size of 224×224. The architecture of ResNet_V2_50 is shown in Figure 5.

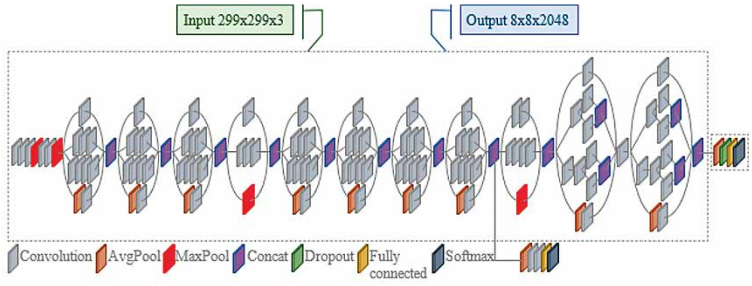

Figure 4 Architecture of inception_V3
Slika 4. Arhitektura modela Inception_V3

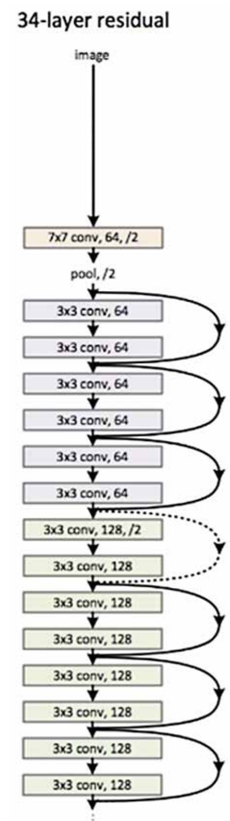

Figure 5 Architecture of ResNet_V2_50 (He *et al*., 2015)
Slika 5. Arhitektura modela ResNet_V2_50 (He *et al*., 2015.)

3.2.2 Fine-tuning
3.2.2. Fino podešavanje

A wide breadth of information can be found in the weights in pre-trained models. This allows for fine-tuning to be carried out and the new model can be trained faster using this information. High achievement can be succeeded in the 2–4 epoch by using learning transfer, even for some problems. The requirement of a large-scale dataset for education purposes is arguably the most significant disadvantage when it comes to models created from scratch, as creating these datasets is very time consuming. The use of pre-trained models and fine-tuning them can give results with less data but still maintain high performance parameters. The simple operation of adding a new fully linked layer(s) to pre-trained models improves success. Fine-tuning, on the other hand, requires the CNN architecture to not only be updated but also re-trained in order to learn new object classes.

In the transfer learning method, the existing layers in the pre-trained models are frozen while the learning is performed; in other words, the weights in these layers should not change. During the training process, random weights associated with the newly added layers are changed. This process is called fine-tuning. Figure 6 illustrates the fine-tuning method in transfer learning.

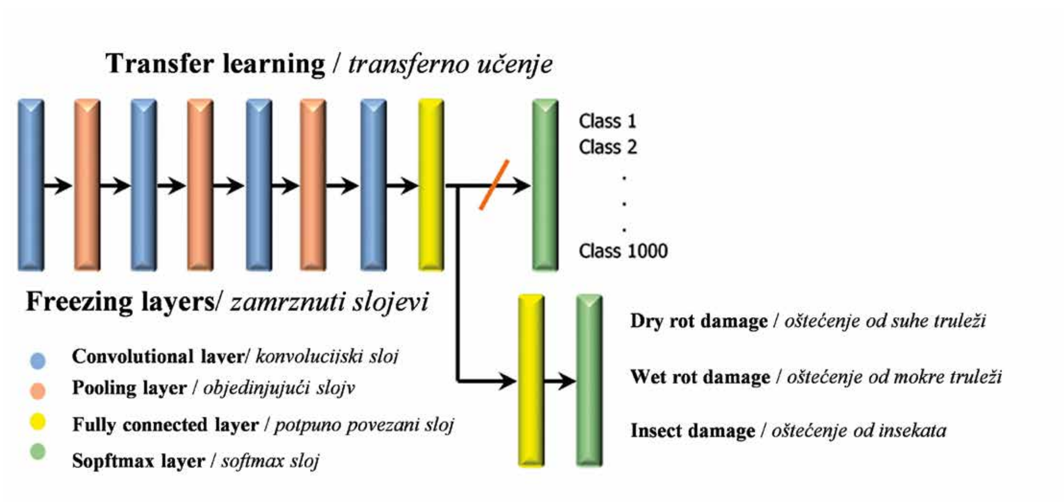

Figure 6 Fine-tuning in transfer learning
Slika 6. Fino podešavanje u transfernom učenju

4 METHODOLOGY
4. METODOLOGIJA

This study aimed to identify damaged areas in traditional wooden structures using a deep classification system based on the classification method. In this study, three types of damage found in wooden structures were considered: dry rot, wet rot, and insect damage. Initially, images depicting these types of damage were collected from various, primarily online, sources. The images were then cropped appropriately and resized to create the dataset used to train the model. A transfer learning technique was used to achieve higher accuracy, instead of training a model from scratch. A selender sample, which is one of the traditional wooden structures found in the region of Trabzon, Turkey, was selected for the experimental study. The analysis of MobileNet_V2, Inception_V3 and ResNet_V2_50 models were performed to determine damaged areas on the serender structure, and results were compared to determine the most reliable and robust pre-trained model. Pre-trained MobileNet_V2, Inception_V3, ResNet_V2_50 are the convolutional neural networks commonly used for an image classification task. The three pre-trained networks were analyzed following the framework as seen in Figure 6. The idea behind MobileNet is to use depthwise separable convolutions to build lighter deep neural networks. Inception was developed with the aim of reducing the computational load of deep neural networks while achieving state-of-the-art performance. While Inception focuses on computational cost, ResNet focuses on computational accuracy and was implemented to increase said accuracy. Due to the phenomena described above, these three networks, which are most commonly used, were deemed the best choices for the experimental study.

Typical input image sizes for MobileNet_V2 , Inception_V3, ResNet_V2_50 are 224×224, 299×299 and 224×224, respectively. A separate set of images not used by the pre-trained model was used for verification to examine the robustness of the models. The three pre-trained neural networks were trained on a data set by using 9140 images; as a result. the algorithms developed can classify new images into 1,000 different object categories. Via deep learning, each network can detect images based on unique features representative of a particular category. By replacing the last fully connected layer, as shown in Figure 7, and retraining (the fine-tuning of deeper layers) the neural network with the new dataset, the neural network is then able to detect specific types of wood damage.

TensorFlow Hub (Abadi *et al.*, 2016) and Keras (Ketkar, 2017) were used in all analyses. Tensorflow is an open source machine learning platform used to perform high-performance numerical calculations. It provides excellent architecture support by enabling easy deployment of computations across a variety of platforms, from desktops to server clusters, mobile devices, and edge devices. The remainder of the imports consists of additional helper functions, followed by NumPy for numerical processing and cv2 for OpenCV bindings.

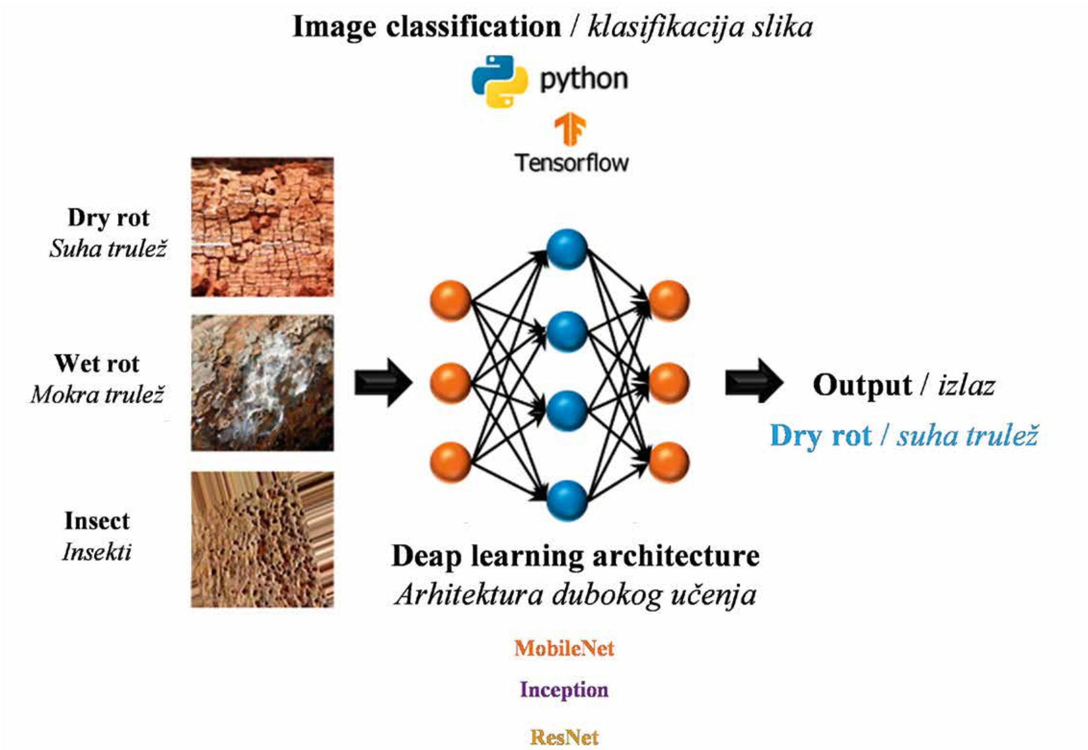

Figure 7 Framework for wood damage detection using pre-trained neural network
Slika 7. Okvir za otkrivanje oštećenja drva uz pomoć unaprijed osposobljene neuronske mreže

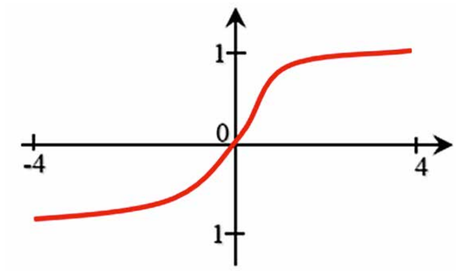

Figure 8 Softmax classification function (Nwankpa *et al*., 2018)
Slika 8. Funkcija klasifikacije *Softmax* (Nwankpa *et al*., 2018.)

There are two types of regularization methods, the L1 and L2, to regulate neural networks. However, the L2 regularization technique is used more often in deep neural networks as it is more advantageous in terms of computational efficiency. The present study used the Softmax classification function (Figure 8), which is used for multiclass classification. The softmax step can be seen as a generalized logistic function that takes a vector of $x \in R^n$ scores as input and creates the probability vector $p \in R^n$ output from the softmax function at the end of the architecture. It is defined as follows:

$$p = \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix} \text{whereas } P_i = \frac{e^{x_i}}{\sum_{j=1}^{n} e^{x_j}}$$

4.1 Dataset
4.1. Skup podataka

Images of different sizes and with a variety of different resolutions were obtained from many sources, including photos taken by mobile phones and copyright-free images obtained from the web, for three types of wood element damage. These images amounted to a total of 9140 pictures, and data augmentation methods were used to increase the dataset size. The data are divided into three main categories: dry rot damage (3245 images), wet rot damage (2877 images) and insect damage (3018 images). For the classification experiments, the dataset was randomly split into 80 % - 20 %, where 80 % was used for training and 20 % was used for testing and validation. The total number of images used as training data amounted to 7312: dry rot damage (2596 images), wet rot damage (2302 images), and insect damage (2414 images). A wide variety of image magnifications have been applied to the training set, including rescaling, rotation, height and width shifts, and horizontal and vertical flips in order to avoid over-fitting and better generalization. The remaining 1828 images of the original 9140 were used as testing data. Samples of the images used are shown in Figure 9.

4.2 Verification of models
4.2. Verifikacija modela

4.2.1 Loss function
4.2.1. Funkcija gubitka

The categorical cross-entropy loss function, which is common in multi-class classification tasks, was used in the study. Cross-entropy generally calculates the difference between two probability distributions for a given random variable/event set.

The categorical cross-entropy loss function calculates the loss of a sample using Eq. 3:

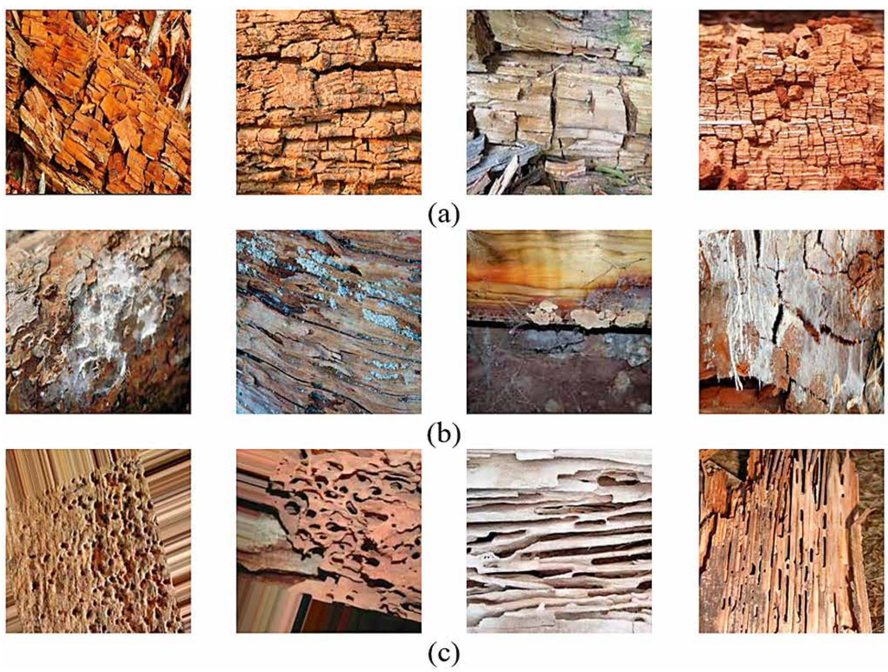

Figure 9 Image samples for (a) dry rot damage, (b) wet rot damage, and (c) insect damage
Slika 9. Uzorci slika za (a) oštećenja od suhe truleži, (b) oštećenja od mokre truleži i (c) oštećenja od insekata

$$Loss = -\sum_{i=1}^{output\ size} y_{\mathrm{i}} \cdot log\ \hat{y}_{\mathrm{i}} \quad (3)$$

Where defines the *i*-th scalar value in the model output, denotes the corresponding target value, and is the number of scalar values in the model output.

The network was trained to more than 30 epochs using a batch size of 32. The sizes of the input images are fixed to height x width = 224 × 224, 299 × 299 and = 224 × 224 for MobileNet_V2, Inception_V3 and ResNet_V2_50 per epoch, respectively. The final accuracy recorded at the end of the 30th epoch for MobileNet_V2, Inception_V3 and ResNet_V2_50 models was 96.38 %, 96.49 %, and 95.39 % for the training, respectively, and with validation values close to the training. The final loss values for MobileNet_V2, Inception_V3 and ResNet_V2_50 models were 0.13, 0.31 and 0.34 for training, respectively, and with validation sets close to the training. In general, the classification ability of loss values below 0.2 is highly acceptable (Fei *et al*., 2020). Accuracy and loss graphs of training and validation for the three network models are given in Figure 10a-c.

The graphs of accuracy for the three models in Figure 10 illustrate the success of the training although the trend for accuracy on both validations and training datasets continued to slightly increase for the last few epochs. It is clear that the models do not over-learn from the training and datasets, and that the validation curves still show a slight increase. However, since the validation curves are adjacent to the training curves and do not diverge from the training curves, none of the models had overfitting problems during training. When these three models are examined, it is evident that MobileNet_V2 is the most robust model in both accuracy and loss of graphic values. In light of this evidence, this model was used for all subsequent analyses.

The models were evaluated to test their robustness, which included a prediction test performed on new images. Upon completion of the test, the accuracy rate for all three models was over 95 %. A sample study of the correct classification of different types of damage images is provided in Figure 11.

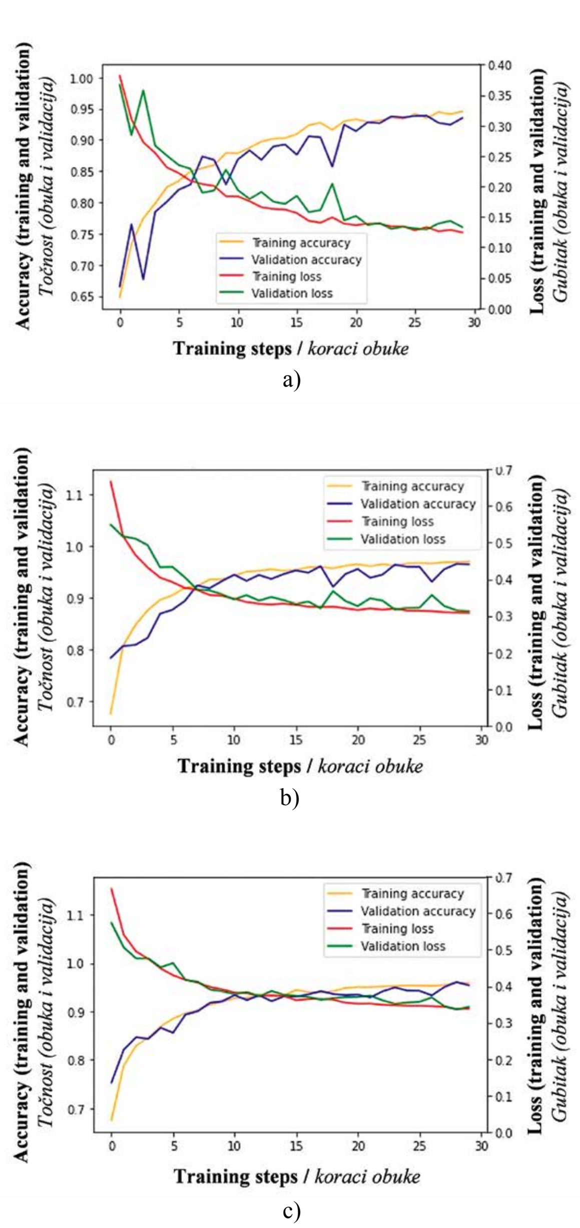

Figure 10 Comparative model accuracy and loss diagram. In sub-figure a) MobileNet_V2, (b) Inception_V3 and (c) ResNet_V2_50 models

Slika 10. Usporedna točnost modela i dijagrami gubitaka: a) model MobileNet_V2, b) model Inception_V3, c) model ResNet_V2_50

5 RESULTS
5. REZULTATI

According to the class of damage previously determined using the image data, all the selected structure regions were photographed to identify the location of damaged areas on a real building model. These critical locations are numbered numerically on the selender photos (Figure 12). The MobileNet_V2 model, identified as the most robust model as previously stated, was used to estimate the class of damage type if found to be present in a given image. As seen in the images in Figure 12, representative images showing different damage types were accurately localized using this method. Figure 12 shows predicted damage in yellow and incorrectly estimated damage in blue.

The pictures obtained from the analyses, the types of damage, and their accuracy rates are labeled on the photo by the computer. As a result of the photos taken, it was observed that these wooden selender structural elements were exposed to all three types of damage that the model was trained on. However, as observed in Figure 12, structural elements were mostly

Figure 11 Dataset used in this study for (a) MobileNet_V2, (b) Inception_V3 and (c) ResNet_V2_50 models
Slika 11. Skup podataka upotrijebljenih u ovoj studiji za modele: a) MobileNet_V2, b) Inception_V3, c) ResNet_V2_50

exposed to insect damage with only one selender exposed to wet rot damge and one selender to dry rot damage. When the predicted accuracy rates were taken into consideration, it was observed that all damage estimates were equal to 100 % and over 99 %.

As seen in Figure 12 (a), where photos obtained for some parts of the sample wooden structure were evaluated, an incorrect classification was made. Insect damage was estimated to exist in all of these classifications. When D11 damage is evaluated with human eyes in Figure 12 (a), it is possible to view this damage as both insect damage and dry rot damage. However, since double or multiple classifications were not made during the training phase, it was estimated that this damage was in fact only insect damage as identified by the model's superior vision. It is important to note that, in the evaluation of D11, 12, D13, and D14 images, there were false predictions, with the model identifying damage when there was none. The issue appears to have been the photo texture, which is similar in appearance to the images of insect damage used in training.

6 DISCUSSION AND SUGGESTIONS FOR FUTURE RESEARCH
6. RASPRAVA I PRIJEDLOZI ZA BUDUĆA ISTRAŽIVANJA

In civil engineering, it is essential to determine the potential damage that can occur in structures, estimate the structure behavior under the influence of external loads, and repair and strengthen said structures with this information in mind. For this reason, in this study, a traditional wooden structure found in Turkey, known as a selender, was chosen as an example, and damage assessment studies based on the deep learning method were carried out. Learning transfer strategies, a state-of-the-art method, have been used to detect damage to wood that occurs in building elements that make wooden structures. The data, classified by damage type, was determined with pre-trained deep convolutional neural network models (MobileNet_V2, Inception_V3, ResNet_V2_50). The deep learning methods used were then compared, and the possibility and type

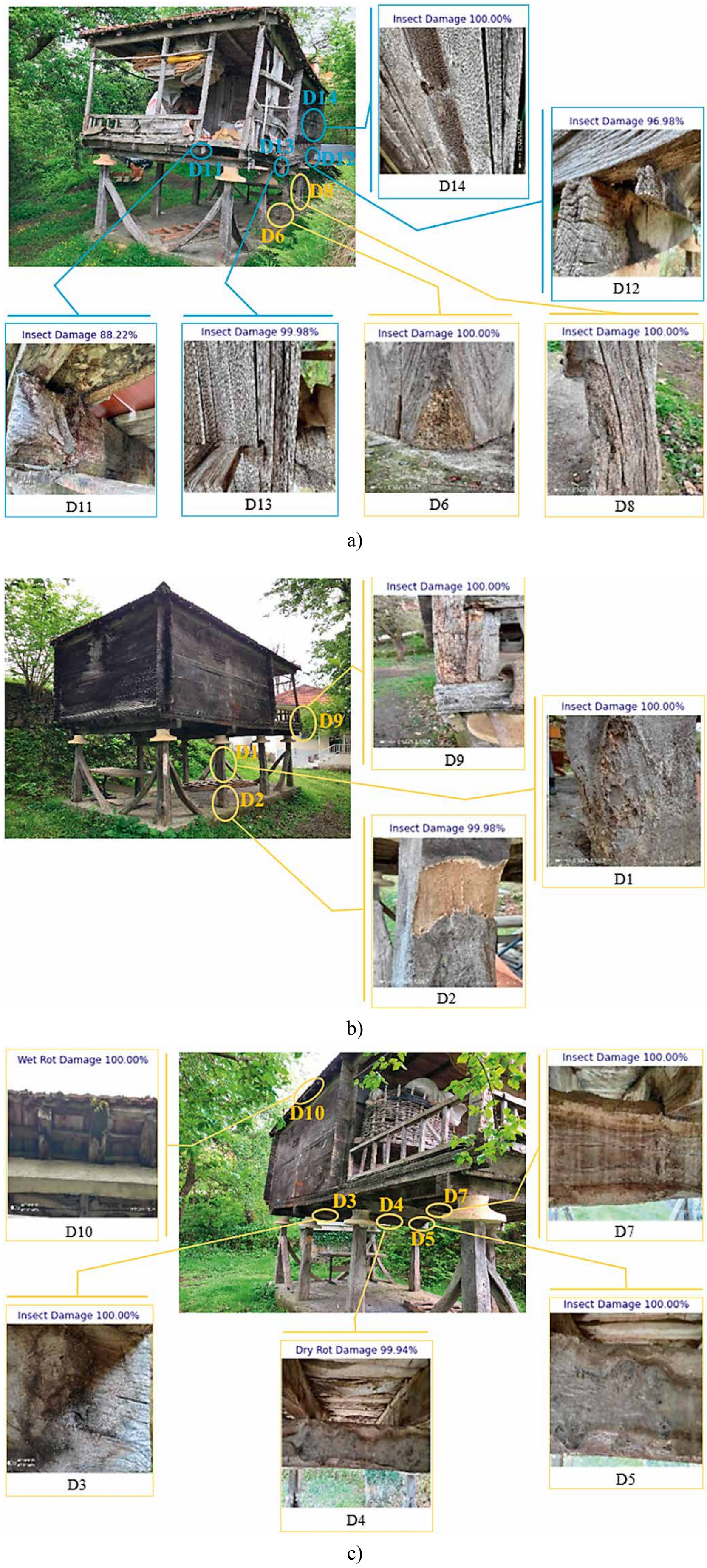

Figure 12 Examples of application images classified as "Insect Damage" classes with MobileNet_V2
Slika 12. Primjeri slika aplikacija klasificiranih kao klase „oštećenja od insekata" u modelu MobileNet_V2

of damage faced by the chosen structure were obtained using the most appropriate method.

Images of different sizes and resolutions were obtained from various sources, including photographs taken with a mobile phone and copyright-free images from the web, as representative of three specific types of wood element damage. These images were increased in number by applying the data augmentation method to the data set. Features obtained using pre-trained models and their labels are divided into training and test sets at 80 % and 20 %. These features are given as input to various classifiers and their labels. Analysis revealed that the lowest loss ratio from these classifiers came from the Multilayer Sensor MobileNet_V2 with 0.13, while the second-lowest loss ratio came from the Inception_V3 classifier with 0.31. The third lowest loss ratio was obtained with the ResNet_V2_50 classifier with 0.34. The accuracy rates of MobileNet_V2, (b) Inception_V3 and (c) ResNet_V2_50 models were identified as 96.38 %, 96.49 %, and 95.39 %, respectively. The accuracy points obtained via these three models are of similar significance. However, when considering the loss ratios, it is apparent that in terms of efficiency and accuracy rates the multi-layer MobileNet_V2 classifier is the most robust.

Three data sets were used in the present study; however, future research can expand on this by using more data sets. Due to the low number of labeled wood element damage data sets available in the literature, it is possible to decrease the ratio of false estimation based on the class by increasing and /or improving the class of wood damage images. The number of images used for training, as well as the correct classification of images, will increase the accuracy of prediction rates of CNN models. The present study identified the damage detection rate of the models used as being over 95 %. This is persuasive evidence of the viability of these models in the classification of types of wood damage.

The ability to accurately identify features of a given object is a direct result of having knowledge and experience of other objects, which varies from person to person. Object recognition in computer science requires the translation of these features into a numerical form. These property values are then processed in a decision making program called a classifier. In this way, each attribute can be associated with a class label. In this way, even a person who does not have any experience with wood damage can easily understand the type of damage, the risk to the structure, and how best to ameliorate the problem.

In future studies, it is considered to identify additional types of damage using various pre-trained models, and to predict which elements in wooden structures cause strength reduction and to what extent. It will also be necessary to take into account the estimation of the dynamic behavior of wooden structures under dynamic loads, such as earthquakes and strong winds, to better calculate potential risks and successfully account for them.

Data availability statement
– Izjava o dostupnosti podataka

Data, models, or codes that support this study's findings are available from the corresponding author upon reasonable request.

5 REFERENCES
5. LITERATURA

1. Abadi, M.; Barham, P.; Chen, J.; Chen, Z.; Davis, A.; Dean, J.; Devin, M.; Ghemawat, S.; Irving, G.; Isard, M.; Kudlur, M.; Levenberg, J.; Monga, R.; Moore, S.; Murray, D. G.; Steiner, B.; Tucker, P.; Vasudevan, V.; Warden, P.; Wicke, M.; Yu, Y.; Zheng, X., 2016: TensorFlow: A System for Large-Scale Machine Learning, Business Opp.
2. Cavalin, P.; Oliveira, L. S.; Koerich, A. L.; Britto, A. S., 2006: Wood defect detection using grayscale images and an optimized feature set. IECON 2006 – 32nd Annual Conference on IEEE Industrial Electronics, 3408-3412. https://doi.org/10.1109/IECON.2006.347618
3. Fei, R.; Yao, Q.; Zhu, Y.; Xu, Q.; Li, A.; Wu, H.; Hu, B., 2020: Deep learning structure for cross-domain sentiment classification based on improved cross entropy and weight. Scientific Programming, 2020, 3810261. https://doi.org/10.1155/2020/3810261
4. Hassaballah, M.; Awad, A. I.; 2020: Deep learning in computer vision: Principles and applications, 1st ed. CRC Press, Qena, Egypt.
5. He, K.; Zhang, X.; Ren, S.; Sun, J., 2015: Deep residual learning for image recognition. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, pp. 770-778. https://doi.org/10.1109/CVPR.2016.90
6. He, T.; Liu, Y.; Xu, C.; Zhou, X.; Hu, Z.; Fan, J., 2019: A fully convolutional neural network for wood defect location and identification. IEEE Access, 7: 123453-123462. https://doi.org/10.1109/ACCESS.2019.2937461
7. He, T.; Liu, Y.; Yu, Y.; Zhao, Q.; Hu, Z., 2020: Application of deep convolutional neural network on feature extraction and detection of wood defects. Measurement, 152, 107357. https://doi.org/10.1016/j.measurement.2019.107357
8. Hu, J.; Song, W.; Zhang, W.; Zhao, Y.; Yilmaz, A., 2019: Deep learning for use in lumber classification tasks. Wood Science and Technology, 53: 505-517. https://doi.org/10.1007/s00226-019-01086-z
9. Jabo, S., 2011: Machine vision for wood defect detection and classification. MSc Thesis, Department of Signals and Systems, Chalmers University of Technology, Göteborg.
10. Jones, S.; Kick-Rack, J.; Pound, W., 2007: Wood-Destroying Insect Diagnostic Inspection. Ohio, USA.
11. Ketkar, N., 2017: Introduction to Keras. In: Deep Learning with Python. Apress, Berkeley, CA. https://doi.org/10.1007/978-1-4842-2766-4_7
12. LeCun, Y.; Bengio, Y.; Hinton, G., 2015: Deep learning. Nature, 521: 436-444. https://doi.org/10.1038/nature14539

13. Nwankpa, C.; Ijomah, W.; Gachagan, A.; Marshall, S., 2018: Activation functions: comparison of trends in practice and research for deep learning. In: Proceedings of 2nd International Conference on Computational and Technology (INCCST 20), MUET Jamshoro.
14. Özgel Felek, S., 2020: Doğu karadeniz yerel mimariye ait serender ve dünya'dan benzer yapılar. Karadeniz İncelemeleri Dergisi, 28: 525-546.
15. Pan, S. J.; Yang, Q., 2010: A survey on transfer learning. IEEE Transactions on Knowledge and Data Engineering, 22 (10): 1345-1359. https://doi.org/10.1109/TKDE.2009.191
16. Robinson, W. H., 2005: Urban insects and arachnids: A handbook of urban entomology. 1st ed. Cambridge University Press. https://doi.org/10.1017/CBO97805115 42718
17. Salman, H.; Grover, J.; Shankar, T., 2018: Hierarchical reinforcement learning for sequencing behaviors. arXiv:1803.01446v1 [cs.RO]. https://doi.org/10.48550/arXiv.1803.01446
18. Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.-C., 2018: MobileNetV2: Inverted residuals and linear bottlenecks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4510-4520.
19. Schmidhuber, J., 2015: Deep learning in neural networks: An overview. Neural Networks, 61: 85-117. https://doi.org/10.1016/j.neunet.2014.09.003
20. Sioma, A., 2015: Assessment of wood surface defects based on 3D image analysis. Wood Research, 60 (3): 339-350.
21. Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A. A., 2017: Inception-v4, inception-ResNet and the impact of residual connections on learning. In: Proceedings of 31st AAAI Conference on Artificial Intelligence, pp. 4278-4284.
22. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A., 2015: Going deeper with convolutions. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 1-9. https://doi.org/10.1109/CVPR.2015.7298594
23. Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z., 2016: Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 2818-2826. https://doi.org/10.1109/CVPR.2016.308
24. Urbonas, A.; Raudonis, V.; Maskeliunas, R.; Damaševičius, R., 2019: Automated identification of wood veneer surface defects using faster region-based convolutional neural network with data augmentation and transfer learning. Applied Sciences, 9, 4898. https://doi.org/10.3390/app9224898
25. van Rossum, G.; Drake, F. L. Jr., 1995: Python reference manual. Centrum voor Wiskunde en Informatica Amsterdam, Computer Science Department of Algorithmics and Architectur, Report, CS-R9525.
26. Wenshu, L.; Lijun, S.; Jinzhuo, W., 2015: Study on wood board defect detection based on artificial neural network. The Open Automation and Control Systems Journal, 7: 290-295. https://doi.org/10.2174/1874444301507010290

Corresponding address:

KEMAL HACIEFENDIOĞLU

Karadeniz Technical University, Department of Civil Engineering, 61080, Trabzon, TURKEY,
e-mail: kemalhaciefendioglu@ktu.edu.tr