UDC 004.93

PREDICTION THE ACCURACY OF IMAGE INPAINTING USING TEXTURE DESCRIPTORS

Kolodochka D. O. – Post-graduate student of the Institute of Computer Systems, Odesa Polytechnic National University, Odesa, Ukraine.

Polyakova M. V. – Dr. Sc., Associate Professor, Professor of the Department of Applied Mathematics and Information Technologies, Odesa Polytechnic National University, Odesa, Ukraine.

Rogachko V. V. – Student of the Institute of Computer Systems, Odesa Polytechnic National University, Odesa, Ukraine.

ABSTRACT

Context. The problem of filling missing image areas with realistic assumptions often arises in the processing of real scenes in computer vision and computer graphics. To inpaint the missing areas in an image, various approaches are applied such as diffusion models, self-attention mechanism, and generative adversarial networks. To restore the real scene images convolutional neural networks are used. Although convolutional neural networks recently achieved significant success in image inpainting, high efficiency is not always provided.

Objective. The paper aims to reduce the time consumption in computer vision and computer graphics systems by accuracy prediction of image inpainting with convolutional neural networks.

Method. The prediction of image inpainting accuracy can be done by an analysis of image statistics without the execution of inpainting itself. Then the time and computer resources on the image inpainting will not be consumed. We have used a peak signal-to-noise ratio and a structural similarity index measure to evaluate an image inpainting accuracy.

Results. It is shown that a prediction can perform well for a wide range of mask sizes and real-scene images collected in the Places2 database. As an example, we concentrated on a particular case of the LaMa network versions although the proposed method can be generalized to other convolutional neural networks as well.

Conclusions. The results obtained by the proposed method show that this type of prediction can be performed with satisfactory accuracy if the dependencies of the SSIM or PSNR versus image homogeneity are used. It should be noted that the structural similarity of the original and inpainted images is better predicted than the error between the corresponding pixels in the original and inpainted images. To further reduce the prediction error, it is possible to apply the regression on several input parameters.

KEYWORDS: image inpainting, accuracy prediction, LaMa network, texture descriptor, co-occurence matrix.

ABBREVIATIONS

CNN is a convolutional neural network;
LaMa is a Large Mask Inpainting;
ReLU is a Rectified Linear Unit;
GLCM is a gray-level co-occurrence matrix;
PSNR is a peak signal-to-noise ratio;
SSIM is a structural similarity index measure;
MSE is a mean squared error;
MAE is a mean absolute error.

NOMENCLATURE

n is a number of image rows;
m is a number of image columns;
(x,y) are coordinates of the image pixel;
$\mathbf{I}(x,y)$ is a vector function representing an image by color channels;
$I_R(x,y)$ is a red channel of an image;
$I_G(x,y)$ is a green channel of an image;
$I_B(x,y)$ is a blue channel of an image;
$M(x,y)$ is a binary mask;
$\circ$ is an element-by-element product of matrices;
f_θ is an inpainting network;
a_i is a coefficient of the polynomial of the 1st degree;
b_i is a coefficient of the polynomial of the 2nd degree;
d_i is a coefficient of the polynomial of the 3rd degree;
q_i is a coefficient of the inverse square root function;
r_i is a coefficient of a logarithmic function;
G is a gray-level co-occurrence matrix;
H is an image entropy;
W is an image homogeneity;
U is an image uniformity;
D_i is a texture descriptor;
P_i is a measure of image inpainting accuracy;
S is the size of the missing image area;
$h_i(\bullet)$ is a function of the dependence of P_i on D_i;
$\mathbf{W}_i$ is a vector of parameters of the function $h_i(\bullet)$;
L is a number of intensity levels in the image;
$I(v, w)$ is a luminance difference between images v, w;
$c(v, w)$ is a contrast difference between images v, w;
$s(v, w)$ is a structure difference between images v, w;
m_v is a local mean of image v;
σ_v is a standard deviation of image v;
σ_{vw} is a cross-covariance for images v and w.
C_i is a positive constant;
α is a positive constant;
β is a positive constant;
γ is a positive constant.

INTRODUCTION

Image inpainting means filling missing image areas with realistic assumptions in computer vision and computer graphics [1, 2]. Often, when photographing, users can encounter unwanted scene elements, for example, random persons or objects that need to be deleted. Before publishing the photo, you may want to make changes to correct the composition. In this case, image inpainting helps to remove unwanted objects and restore the image. Another case of application is the restoration of old photos that have been physically damaged.

DOI 10.15588/1607-3274-2025-2-5

The missing areas in an image can be inpainted by various approaches such as diffusion models (face and expressions inpainting [3, 4]) and self-attention mechanism (object removal in remote sensing images [2, 5, 6]). Generative adversarial networks are used for general image object removal and image desensitization which replaces sensitive information in images [2, 7]. To restore modern life and industrial images single- and multi-subnet CNNs are applied [8–10].

When filling an image area, it is necessary to select an inpainting method depending on the size of the missing area and the properties of the image. The CNNs known from the literature do not always provide high efficiency [1, 2, 8–10]. The question arises about the advisability of using a particular CNN and, accordingly, the consumption the time and computer resources on the image inpainting. Therefore, it is desirable to predict the accuracy of the filling of an image area of a certain size. It is supposed that the selected CNN is applied to the specific type of researched images.

The object of study is inpainting of real scene images in computer vision and computer graphics systems.

The subject of the study is methods of accuracy prediction of image inpainting using CNNs.

The paper aims to reduce the time consumption in computer vision and computer graphics systems by accuracy prediction of image inpainting with CNNs.

1 PROBLEM STATEMENT

The three-channels real scene image is defined as $\mathbf{I}(x,y)=(I_R(x,y), I_G(x,y), I_B(x,y))$, where $x=1, \ldots, n$; $y=1, \ldots, m$. To represent the missing areas of the image a mask is introduced. It is a binary image $M(x, y)$ of the same size as each channel of the original image. The mask is element-by-element multiplied by image channels, and the image with missing areas is represented as $\mathbf{I_M}(x,y)=(I_R(x,y)\circ M(x, y), I_G(x,y)\circ M(x, y), I_B(x,y)\circ M(x, y))$ [8–9, 11].

Let us suppose that the CNN $f_\theta(\bullet)$ with parameters θ was preliminarily trained to inpaint the images. It outputs the image $\mathbf{I_{in}}(x,y) = f_\theta(\mathbf{I_M}(x,y))$ which approximates the original image $\mathbf{I}(x,y)$ in the sense of some criterion [8].

To predict the accuracy of the image inpainting by the considered network, an image feature (one or more) is selected. To evaluate this feature, descriptors $D_1, D_2, ..., D_k$ are formed. They will be used as independent variables representing the input information for the prediction. Besides, the prediction can be significantly influenced by additional factors. In our case, this is the size of the missing image area S.

Next, the output variables are selected. These are measures evaluating the accuracy of the image inpainting by the selected network: $P_1, P_2, ..., P_l$. It is necessary to define the dependence of measures $P_1, P_2, ..., P_l$ on descriptors $D_1, D_2, ..., D_k$ and factor S :

$$P_i=h_1(D_1, \mathbf{W}_1, S),\ P_i=h_2(D_2, \mathbf{W}_2, S),\ \ldots,\ P_i=h_k(D_k, \mathbf{W}_k, S),\ \ i=1, \ldots, l; \tag{1}$$

and to estimate vectors of parameters $\mathbf{W}_1, \mathbf{W}_2, ..., \mathbf{W}_k$ of these dependences. Further for each P_i, $i=1, ..., l$, we determine from a set of dependencies (1) the dependency $P_i=h_{j(i)}(D_{j(i)}, \mathbf{W}_{j(i)}, S)$, $j(i)\in\{1, ..., k\}$, with the highest value of the selected measure of approximation accuracy. Interpolation or extrapolation of the functions, $P_i=h_{j(i)}(D_{j(i)}, \mathbf{W}_{j(i)}, S)$, $i=1, ..., l$; will predict the accuracy of image inpainting by the network $f_\theta(\bullet)$.

In this paper we propose a method to predict the accuracy of the real scene image inpainting with the selected neural network. To obtain the solution to this problem we determined the factors that influenced the inpainting accuracy measures. Based on these factors we selected the predictors and determined the structure of dependence between the predictors and the measures of inpainting accuracy. After the estimation of the parameters of such dependence and the approximation accuracy evaluation, we can calculate an input parameter and predict the output parameter based on the dependence existing between them. Having this prediction, one can decide by a human or automatically using certain rules whether to apply considered CNN for image inpainting.

To research the proposed method experimentally we used the selected LaMa convolutional network and varied the size of the missing area which must be inpainted. The size of the missing area defines the complexity of image inpainting and determines the effectiveness of the prediction for the selected network. We also evaluated the accuracy of prediction for the different spectral transforms consisted the LaMa network and for different image inpainting measures. The experimental results presented in this article were obtained using the Places365 dataset [12].

2 REVIEW OF THE LITERATURE

In the literature the ability to predict the CNN efficiency was researched for classification problems in [13]. The prediction of the image denoising efficiency based on the discrete cosine transform was considered in [14–16]. The paper [17] is devoted to the prediction of the dynamics of the signals describing a signature using the CNN. It is needed to detect potential forgeries by identity verification systems. However, the network efficiency has not been predicted for the problem of image inpainting. The recommendations for the selection of image inpainting methods are mostly qualitative. The evaluation of the accuracy of the filling of missing image areas is provided after image inpainting which is time and resource-consuming.

The existing image inpainting approaches are influenced by the content of input images, missing area size, and the ill-posedness of image inpainting problem. The deep learning-based methods in image inpainting allow for improving results by capturing image features and semantic on different scales [18]. Despite the similarity of deep learning methods, they differ in inpainting approaches, network architecture, loss function, etc. To construct the model for inpainting

DOI 10.15588/1607-3274-2025-2-5

accuracy prediction it is advisable to take into account the following four image inpainting approaches [1].

The progressive inpainting fills missing image areas step by step. Specifically, coarse-to-fine, low-to-high-resolution, and structure-to-content inpainting are used. This approach supposes that available information is not sufficient to reconstruct all missing pixels in one step [5, 18].

The structural information-guided inpainting is based on the structure of the known regions, such as segmentation results, edges, depth maps, gradient of color or intensity, etc. These auxiliary cues facilitate the recovery of sharp details and fine structure of the missing areas [19–21].

The convolutions-aware inpainting applies different convolution operators for the generation of missing pixels. For example, the traditional convolution operates by valid pixels as well as by substitute values in the missing areas, which leads to color discrepancy and blurring [22]. To avoid the disadvantages, partial convolution, gated convolution, and bidirectional convolution are used [22–24].

The attention-based inpainting uses the image content from distant spatial locations when CNNs are ineffective for the filling of missing areas. The additional information can be obtained from contextual patches, feature maps, searching for the most similar patch, or the modeling of the underlying distribution of reconstructed images [6, 7, 25].

Let us review these approaches in light of the possibility of their applying to the prediction of image inpainting accuracy. So, progressive inpainting is not well suited for inpainting accuracy prediction due to its generality. Almost all the CNNs based on the convolutions-aware approach due to their architecture. If we will use the same approach to predict the accuracy of the image inpainting by the CNN we can obtain the overtrained prediction model. The attention-based approach uses local patterns of the original image and similar images. It is not clear how these local patterns can influence image inpainting in general and how to elaborate a predictive model based on such a representation. However, the attention-based approach can be used to obtain additional information for image inpainting accuracy prediction, specifically, the properties of the distribution of image colors or intensities.

It should be noticed that the image content and fine structure influence the image inpainting accuracy. Therefore, it is advisable to predict the image inpainting accuracy based on a structural information-guided approach which allows to extraction of the image features and to estimate the feature descriptors.

3 MATERIALS AND METHODS

The proposed method of predicting the accuracy ofThe proposed method of predicting the accuracy of the image inpainting by CNN applies the regression on initial image descriptors. The general idea of the proposed method can be described by the following steps.

Step 1. Preparing the data.

The datasets for image inpainting are selected. The range of sizes and the form of the image missing areas are determined. Then the binary masks modeling the image's missing areas are formed and superimposed on images.

Step 2. The tuning of the CNN.

As a result of the literature review, we choose CNN for image inpainting. If a pre-trained network is used then the parameters of such network are loaded. Otherwise, the training and test sets of images are formed from the selected datasets. The researched network is trained to inpaint the missing areas modeled with binary masks on the images of the training set.

Step 3. The determining of the prediction model variables.

To obtain the output variables of a prediction model, the measures of image inpainting accuracy are selected. The trained network is applied to test set images with missing areas. The inpainting accuracy is evaluated by the selected measures depending on the size of the missing areas of the image.

Next, we choose the image descriptors that estimate the image features influencing the inpainting accuracy. They will be used as input variables of the model. Values of these descriptors are calculated based on original images from the test set.

Step 4. The prediction model selection and model parameters estimation.

The regression model that connects output and input variables is constructed using the fitting of the curves to data points in scatter plots. While scatter plots are created we take into account that original images may have different properties or content of missing areas differs from one image to the other. The regression model parameters are estimated on a training set of images by the least squares approach.

Step 5. The prediction model evaluation and the image inpainting accuracy prediction.

To evaluate the prediction model performance the accuracy of curve fitting is estimated based on the values of input and output variables obtained on the test set images. The regression interpolation or extrapolation predicts the values of the image inpainting accuracy measures relying on the values of descriptors. The obtained results are analyzed to conclude the appropriateness of the considered CNN to inpaint the specific image.

In this paper, we provide prerequisites for decision-making based on the specifics of image inpainting methods. In the following research, we concentrate on a particular case of versions of the LaMa network although the proposed method can be generalized to other CNNs as well.

The LaMa network is considered here to inpaint the images for the following reasons.

(1) This CNN has a simpler architecture than many of the other state-of-the-art networks for image inpainting. In particular, the LaMa network consists of one subnet rather than an ensemble of subnets and has fewer parameters.

DOI 10.15588/1607-3274-2025-2-5

(2) It can implement data processing with fast Fourier transform or discrete wavelet transform leading to several benefits:

a) a higher speed of image processing and network training;

b) successful inpainting of large missing areas of spectral textures;

c) high-quality inpainting of fine details of images and edges of objects.

Let us describe the architecture of the LaMa network (Fig. 1) [8, 9, 11]. This network has inputted an image with an overlayed mask denoting the pixels that need to be inpainted. At first, the initial image is downscaled by a factor of 3. Then the local and global textures are extracted from the obtained image. These global and local features are further passed through convolution layers. Then spectral transform block is additionally applied to global texture channels. The fast Fourier transform or the discrete wavelet transform is implemented in this block. The features of using these transforms in the LaMa network are described and studied in detail in [8, 9, 11]. The outputs of the convolution layers are added up crosswise. Then after batch normalization and ReLU activation, the results of local and global texture processing are concatenated. The described layers excluded from the downscale are repeated and added up with the downscaled image. The LaMa network module embedded between downscaling and upscaling is also repeated nine times. After that, the image is upscaled to its initial size and outputted [8, 9].

Despite the relatively simple LaMa network architecture, the learning of this network is time-consuming. The image impainting by this trained network also requires a lot of time. For example, in the Google Colab environment with a pre-configured NVIDIA Tesla T4 GPU, which has 16 GB of GDDR6 memory and 2,560 CUDA cores it took us 2 to 7 seconds to inpaint an image of size 1024x1024 pixels. At the same time although the LaMa network has demonstrated a high inpainting accuracy, some of the restored images are of insufficient quality. In particular, this is the case of the inpainting of statistical textures and detail-rich images. In [11] it was noted that the quality of the images inpainted with the LaMa network is influenced by the initial image textures. To avoid wasting time and resources we propose to previously evaluate the accuracy of image inpainting by the LaMa network.

Let us define the output variables of a prediction model. We propose to use PSNR and SSIM which estimate the inpainting accuracy in terms of edge quality and structural integrity [26]. Let us suppose that the lowest image intensity level is equal to zero. Then PSNR is evaluated as the ratio between the highest squared intensity $(L-1)$ of the initial image and MSE between the initial and inpainted images [26]:

$$PSNR = 10\times\log_{10}((L-1)^2/MSE).$$

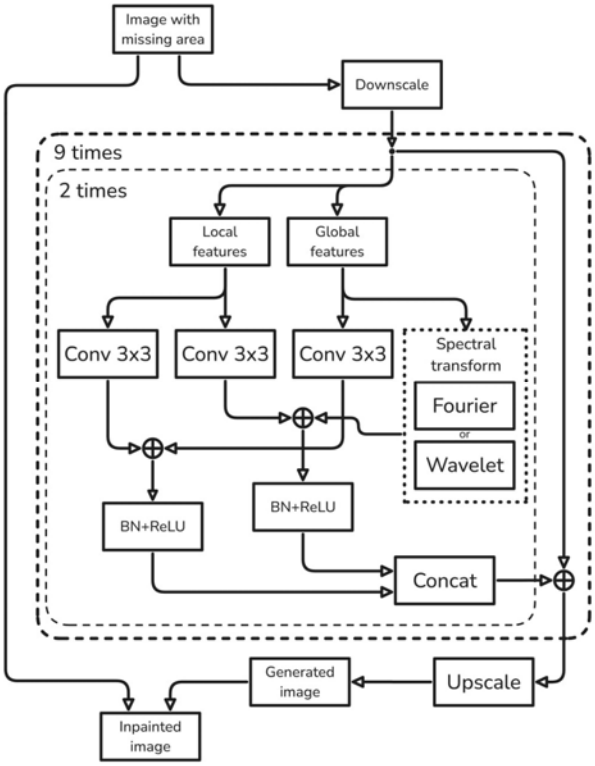


Figure 1 – LaMa network architecture [8, 9]

PSNR compares the edges and fine details in the original and inpainted image and determines differences between them at the pixel level.

Another considered measure is SSIM evaluates the inpainted image in terms of restoring the natural appearance of textures and edges. It estimates the differences in texture, contrast, and structure of the initial and reconstructed images [26]:

$$SSIM(v, w)=l\,(v, w)^{\alpha}\;c\,(v, w)^{\beta}\;s\,(v, w)^{\gamma},$$
$$l\,(v, w) = (2m_v\, m_w + C_1)/(m_v^{\,2} + m_w^{\,2}+C_1),$$
$$c\,(v, w) = (2\sigma_v\,\sigma_w + C_2)/(\sigma_v^{\,2} + \sigma_w^{\,2}+ C_2),$$
$$s\,(x, y)=(\sigma_{vw} + C_3)/(\sigma_v\,\sigma_w + C_3).$$

We suppose that $\alpha=\beta=\gamma=1$, then the SSIM values range between 0 to 1.

Selecting the input variables of the prediction model we observed that the accuracy of image inpainting weakly depends on the texture measures based on the color or intensity histograms. They ignore the spatial relations between pixels, which is important when a texture is described. To avoid the shortcoming it can be taken into consideration not only the distribution of intensities but also the relative positions of pixels in an image. In this way in [27] an operator Q is defined which evaluates the relation between the intensities of two pixels. Based on this operator a GLCM G is determined for a gray-scale image $I(x, y)$ with L possible intensity levels. Each element g_{ij} of the matrix G is the number of times that the pixel pair with intensities l_i and l_j is found in the image in the relation Q, where $1 \le i, j \le L$.

DOI 10.15588/1607-3274-2025-2-5

Based on a GLCM the texture descriptors are introduced in [27]. When the LaMa network efficiency was tested, we observed that the texture descriptors such as uniformity, homogeneity, and entropy changed for different values of the PSNR and SSIM [11]. Therefore, we can to use these texture descriptors as the input variables of the prediction model because their computation is time-saving compared to the image inpainting by CNN.

The uniformity determines the pixel intensity randomness and takes values from the range [0, 1]:

$$U = \frac{1}{n_Q^2} \sum_{i=1}^{L} \sum_{j=1}^{L} g_{ij}^2 ,$$

where n_Q is the number of pixel pairs in the relation Q. Uniformity increases as the square of the probability values, so the less random an image is, the higher its uniformity. The uniformity is equal to 1 for a constant image.

Homogeneity measures the concentration of GLCM element values near the main diagonal by expression

$$W = \frac{1}{n_Q} \sum_{i=1}^{L} \sum_{j=1}^{L} \frac{g_{ij}}{1+|i-j|} .$$

The value of the denominator $(1 + |i - j|)$ decreases as the values of i and j get closer, i.e., as they approach the main diagonal. The range of homogeneity values is [0, 1], with the maximum being achieved when G is a diagonal matrix. The GLCM with the highest values of elements near the main diagonal will correspond to images with a variety of gray-level content and areas of slowly varying intensity values.

Entropy measures the randomness of the elements of GLCM which, in turn, is determined by the randomness of the initial image:

$$H = -\sum_{i=1}^{L} \sum_{j=1}^{L} \frac{g_{ij}}{n_Q} \log_2 \frac{g_{ij}}{n_Q} .$$

The highest value $2\times\log_2 L$ is achieved for matrix GLCM, obtained from an image that is formed by uniform noise then all image intensities are approximately equally probable. The entropy is equal to zero for a constant-intensity image.

4 EXPERIMENTS

Let us consider the construction of the scatter plots for the image inpainting accuracy prediction. To obtain the regression function that describes the dependence between output and input variables, the curve fitting of scatter plots is applied. To take into consideration the different properties of inpainted images the simulation of missing areas of the different sizes is needed. That is why we have formed three separate categories of masks, modeling different levels of image inpainting complexity. These masks randomly uniformly cover 25%, 50%, and 75% of the image area, and are named as narrow, medium, and large, respectively.

We constructed masks of 1–5 straight lines with a slope from 0 to 2π, 1–100 pixels wide, and 10–200 pixels long. As an alternative, masks of 1–4 rectangles with sides of 30–150 pixels were generated with a probability of 0.5. Then the test set included 2,000 images from the Places2 dataset [12]. The missing areas of these images covered by generated masks were reconstructed by the LaMa network. As spectral transform fast Fourier transform or discrete wavelet transform were included in LaMa network architecture. The LaMa-Fourier and LaMa-Wavelet networks are obtained [8, 9]. The results of image inpainting were evaluated with PSNR and SSIM.

Further, the GLCM has been determined and the image uniformity, homogeneity, and entropy have been calculated based on the obtained GLCM for each test image. Next, one mask from each category was generated for each image. After element-by-element multiplication of images on masks, the inpainting was performed using LaMa-Fourier and LaMa-Wavelet networks. In Figure 2 the scatter plots SSIM and PSNR versus uniformity, homogeneity and entropy are shown.

As a basis for prediction, the data points in the obtained scatter plots indicate the main trends. Specifically, the PSNR and SSIM decrease if entropy increases and uniformity and homogeneity decrease. Data for different mask categories are shown in the different scatter plots. The results related to the image inpainting by the LaMa-Wavelet and LaMa-Fourier networks denoted with different symbols.

Next, we consider the estimation of the parameters of the image inpainting accuracy prediction model. The following functions for the scatter-plot fitting were selected.

1. Polynomials of the first, second and third degree $y(x)=a_1x+a_0$, $y(x)=b_2x^2+b_1x+b_0$, $y(x)=d_3x^3+d_2x^2+d_1x+d_0$.
2. Inverse square root function $y(x) = q_1x^{-1/2} + q_0$.
3. Logarithmic function $y(x) = r_1 \ln x + r_0$.

The parameters of the mentioned functions were estimated by the fitting of the scatter plots using the least squares approach. To evaluate the curve fitting results we have selected a curve fitting accuracy measure. This is the R-squared (R^2) value which is estimated as the proportion of the variance in the dependent variable explained by the independent variable in the considered model. The R-squared values lie between 0 and 1, where higher R^2 relates to better curve fitting. To estimate the approximation error the mean squared error (MSE) is used. It estimates the average squared difference between the actual and predicted values of the dependent variable. Lower MSE indicates that the selected model better approximates the actual values. MSE is bounded below by zero and has no higher limit.

DOI 10.15588/1607-3274-2025-2-5

5 RESULTS

The obtained values of R-squared are shown in Tables 1nand 2. We have discarded the low-valued outliers of entropy (2.5% for narrow and medium masks, 5% for large masks), and the high-valued outliers of uniformity (10% for narrow masks, 15% for medium masks, 20% for large masks) to increase the approximation accuracy. As a result, the R-squared values were increased by an average of 7–8%. The best R-squared values were mostly obtained on polynomials of the third order but the other functions used for scatter-plot fitting in the paper have shown similar results.

In Table 3 the MSEs related to SSIM and corresponding to the R-squared from Table 1 are presented without the brackets. The MSEs related to PSNR and corresponding to the R-squared from Table 2 are shown in brackets.

Table 1 – The R-squared for polynomial regression degree 3, the dependent variable is SSIM

Input variable	Narrow masks	Medium masks	Large masks
	LaMa-Fourier		
Entropy	0.3950	0.4192	0.4241
Homogeneity	0.7007	0.7098	0.6920
Uniformity	0.4408	0.4641	0.4500
	LaMa-Wavelet		
Entropy	0.4151	0.4277	0.4257
Homogeneity	0.7203	0.7267	0.7090
Uniformity	0.4541	0.4630	0.4489

The obtained dependencies of PSNR versus homogeneity W for narrow masks are expressed for LaMa-Fourier and LaMa-Wavelet as

$$PSNR=24.3653+14.9912W–2.4763W^2+ 4.7581W^3;$$
$$PSNR=24.1664+18.0167W–11.4362W^2 +13.5988W^3.$$

The SSIM versus homogeneity W for narrow masks is expressed for LaMa-Fourier and LaMa-Wavelet as

$$SSIM=0.8228+0.4694W–0.7282W^2+ 0.4640W^3;$$
$$SSIM=0.8199+0.4644W–0.7077W^2+ 0.4622W^3;$$

for medium masks as

$$SSIM=0.6341+0.8195W–1.0334W^2+0.6318W^3;$$
$$SSIM=0.6291+0.8284W–1.0622W^2+0.6838W^3;$$

for large masks as

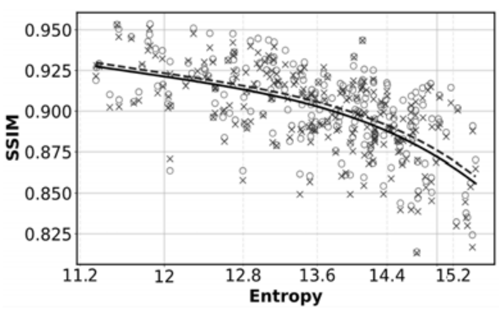
$$SSIM=0.4112+1.1130W–1.1985W^2+0.7524W^3;$$
$$SSIM=0.3988+1.1170W–1.1921W^2+0.7549W^3.$$

Table 2 – The R-squared for polynomial regression, degree 3, the dependent variable is PSNR

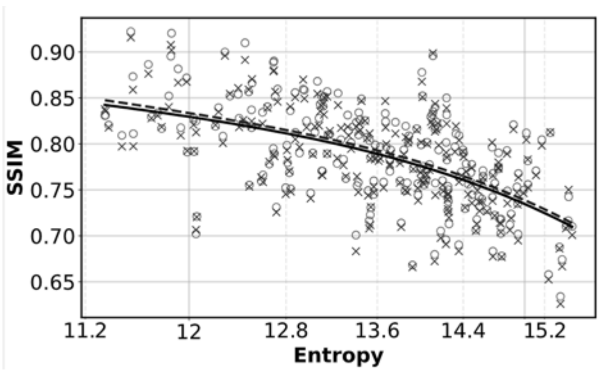
Input variable	Narrow masks	Medium masks	Large masks
	LaMa-Fourier		
Entropy	0.3220	0.2707	0.1457
Homogeneity	0.4826	0.3496	0.1960
Uniformity	0.2720	0.1691	0.0600
	LaMa-Wavelet		
Entropy	0.3329	0.2691	0.1466
Homogeneity	0.5034	0.3523	0.1961
Uniformity	0.2831	0.1620	0.0603

Table 3 – The MSE for polynomial regression, degree 3, the dependent variable is SSIM (PSNR)

Input variable	Narrow masks	Medium masks	Large masks
	LaMa-Fourier		
Entropy	0.0004 (5.0254)	0.0019 (5.6551)	0.0048 (4.8962)
Homogeneity	0.0002 (4.0205)	0.0009 (5.0474)	0.0025 (4.9400)
Uniformity	0.0003 (5.5083)	0.0016 (5.6171)	0.0042 (4.7120)
	LaMa-Wavelet		
Entropy	0.0004 (5.0083)	0.0019 (5.7219)	0.0048 (4.7160)
Homogeneity	0.0002 (3.9224)	0.0009 (5.0739)	0.0025 (4.4509)
Uniformity	0.0004 (4.7620)	0.0017 (6.1228)	0.0043 (4.7074)

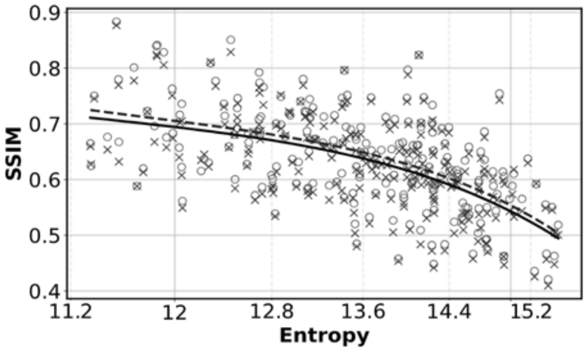


a

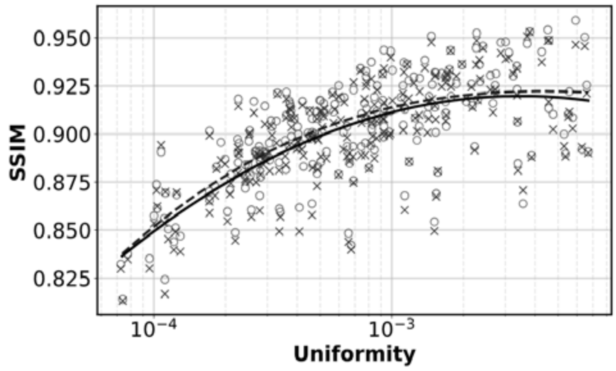


b

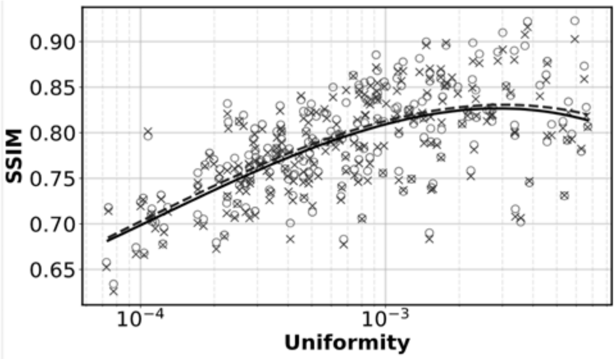
DOI 10.15588/1607-3274-2025-2-5



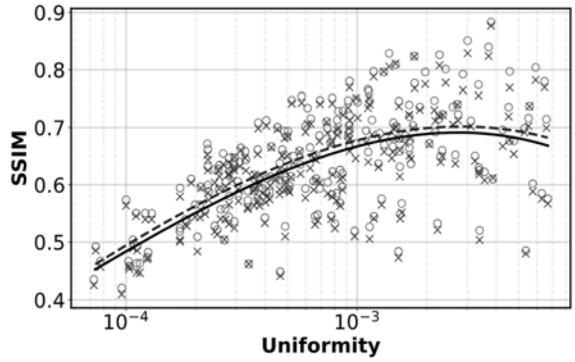
c



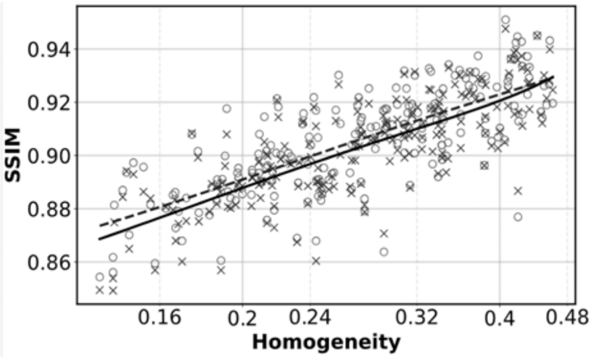
d



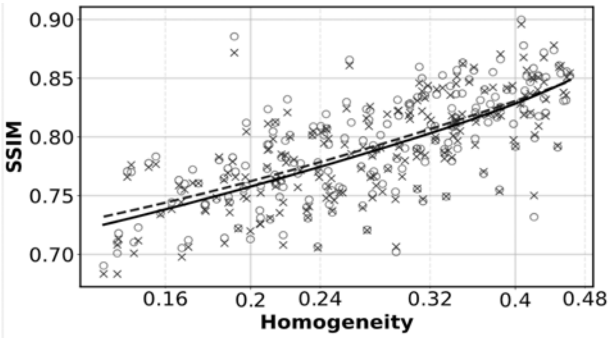
e



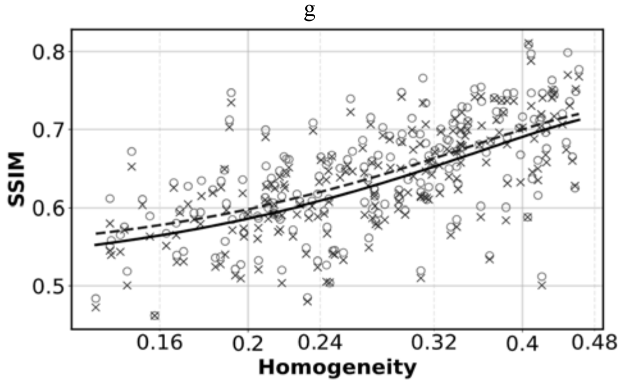
f



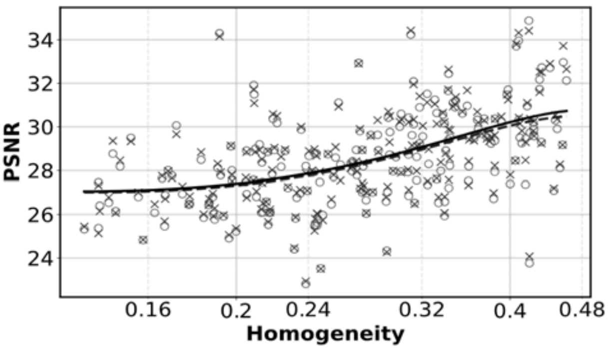
g



h



i



j

Figure 2 – The scatter-plots of SSIM vs entropy: a – narrow mask; b – medium mask; c – large mask; SSIM vs uniformity: d – narrow mask; e – medium mask; f – large mask; SSIM vs homogeneity: g – narrow mask; h – medium mask; i – large mask; PSNR vs homogeneity: j – narrow mask. The data points and the line of polynomial regression of 3rd degree related to LaMa-Fourier results are marked with a circle and dash line; the same objects related to LaMa-Wavelet are marked with a cross and solid line

DOI 10.15588/1607-3274-2025-2-5

6 DISCUSSIONS

Let us consider the dependencies of the SSIM from the texture descriptors if the LaMa-Fourier network is used. The best fitting is achieved for SSIM versus homogeneity (Table 1). Specifically, the approximation by a polynomial of degree 3 gives R^2 from 0.6920 to 0.7098 for different mask size. This indicates that considered model fits the data well. Others texture descriptors worse explain the dependent variable SSIM (R^2 varies from 0.3950 to 0.4241 for entropy and from 0.4508 to 4641 for uniformity). As one can see, the results do not appear to be affected by the mask size.

Now we analyze the dependences the PSNR from the texture descriptors if the LaMa-Fourier network is used (Table 2). The best fitting is again achieved for homogeneity versus PSNR. Specifically, the approximation by a polynomial of degree 3 gives R^2 from 0.1960 to 0.4826 for different mask size. Others texture descriptors even worse explain the PSNR (R^2 varies from 0.1457 to 0.3220 for entropy and from 0.0603 to 0.2831 for uniformity). Therefore, the PSNR can be only predicted from homogeneity and for narrow masks. As for the remaining scatter plots, the proportion of the PSNR variance that is explained by texture descriptors is very low to predict the actual PSNR values.

If the LaMa-Wavelet network is applied then the R-squared values are increased by 3–5% for considered dependencies and categories of masks (Tables 1, and 2). Therefore, the LaMa-Wavelet network is more acceptable for the prediction of the image inpainting accuracy.

The scatter plots for different mask sizes are presented in different figures (Figure 2). It can be observed that the LaMa network versions in general show similar results, i.e. the data points related to the LaMa-Fourier and LaMa-Wavelet do not create separate clusters. Moreover, the scatter plots of dependences of the texture descriptors on the SSIM obtained for the different mask sizes are also similar. However, in the scatter-plots of dependences of the texture descriptors on the PSNR, the compactness of data points is enlarged as the mask size is increased. The cluster of data points is “pressed” to the texture descriptor axis showing the significant lowering of the PSNR of inpainted images as the missing area size increases. It means that the accuracy of image detail in inpainting is decreased as mask size is increased.

The results of the prediction of the image inpainting accuracy on a test set of images from the Places2 dataset are presented in Table 4. MAE was estimated between the actual and predicted values of PSNR and SSIM. It should be noted that the lowest MAE of the prediction we have obtained for narrow mask inpainting. However, the highest MAE is obtained for medium masks. This fact can be explained by a high variety of medium-sized image details.

As an example, we have considered the results of the prediction of the inpainting accuracy for the images from Figure 3. The images from Figures 3, a, b are processed with narrow masks; the images from Figures 3, c, d are inpainted with medium and large masks, respectively. The PSNR was only predicted for images from Figure 3, a, b. The actual PSNR were 24.7131 dB and 30.6880 dB, the predicted PSNR were 30.5912 dB and 31.7558 dB. The SSIM was predicted for images from Figure 3, a–d. The actual values were 0.8926, 0.9281, 0.7819, and 0.5452; the predicted values were 0.9260, 0.9330, 0.8586, 0.6185; respectively.

Table 4 – The MAE of image inpainting accuracy prediction

Narrow masks, PSNR, dB	Narrow masks, SSIM	Medium masks, SSIM	Large masks, SSIM
LaMa-Fourier			
3.7359	0.0262	0.0960	0.0504
LaMa-Wavelet			
3.7443	0.0283	0.0966	0.0705

Finally, we note that computing the texture descriptors, specifically, entropy, homogeneity, and uniformity, took an average of 0.1139 sec per image of 1024x1024 pixels. The PSNR and SSIM calculation takes an average of 0.1605 sec per image of the same size. By comparison, we can mention that image inpainting by the LaMa-Wavelet network took an average of 6.6 sec per image. However, the network’s computations are predominantly GPU-accelerated, making heavy use of hardware-level parallelism, while the accuracy metrics evaluation is done on the CPU (Intel(R) Xeon(R) CPU 2.00GHz, total RAM 12.67 GB) relying on synchronous operations. It should be noted that because of inherent differences in the hardware architecture, the time of the computing of texture descriptors and calculations in the network should not be directly compared.

CONCLUSIONS

The scientific novelty is the proposed method of the prediction of the image inpainting accuracy. The method is based on the set of texture descriptors estimated using the gray-level co-occurrence matrix to predict the values of image inpainting accuracy measures, specifically, the PSNR and SSIM.

The practical significance of the research is in the results obtained for real scene images from the Places2 dataset with the LaMa network applied. These results show that the prediction of the image inpainting accuracy can be performed with satisfactory accuracy if the dependencies of the SSIM or PSNR versus homogeneity are used. The other considered texture descriptors such as entropy and uniformity can be only used to support the prediction. It should be noted that the structural similarity of the original and inpainted images is better predicted than the error between the corresponding pixels in the original and inpainted images. The better approximation accuracy was achieved with the polynomial of 3rd degree if data outliers have been removed. Better prediction accuracy was obtained if the missing areas could be modeled by a narrow mask.

DOI 10.15588/1607-3274-2025-2-5

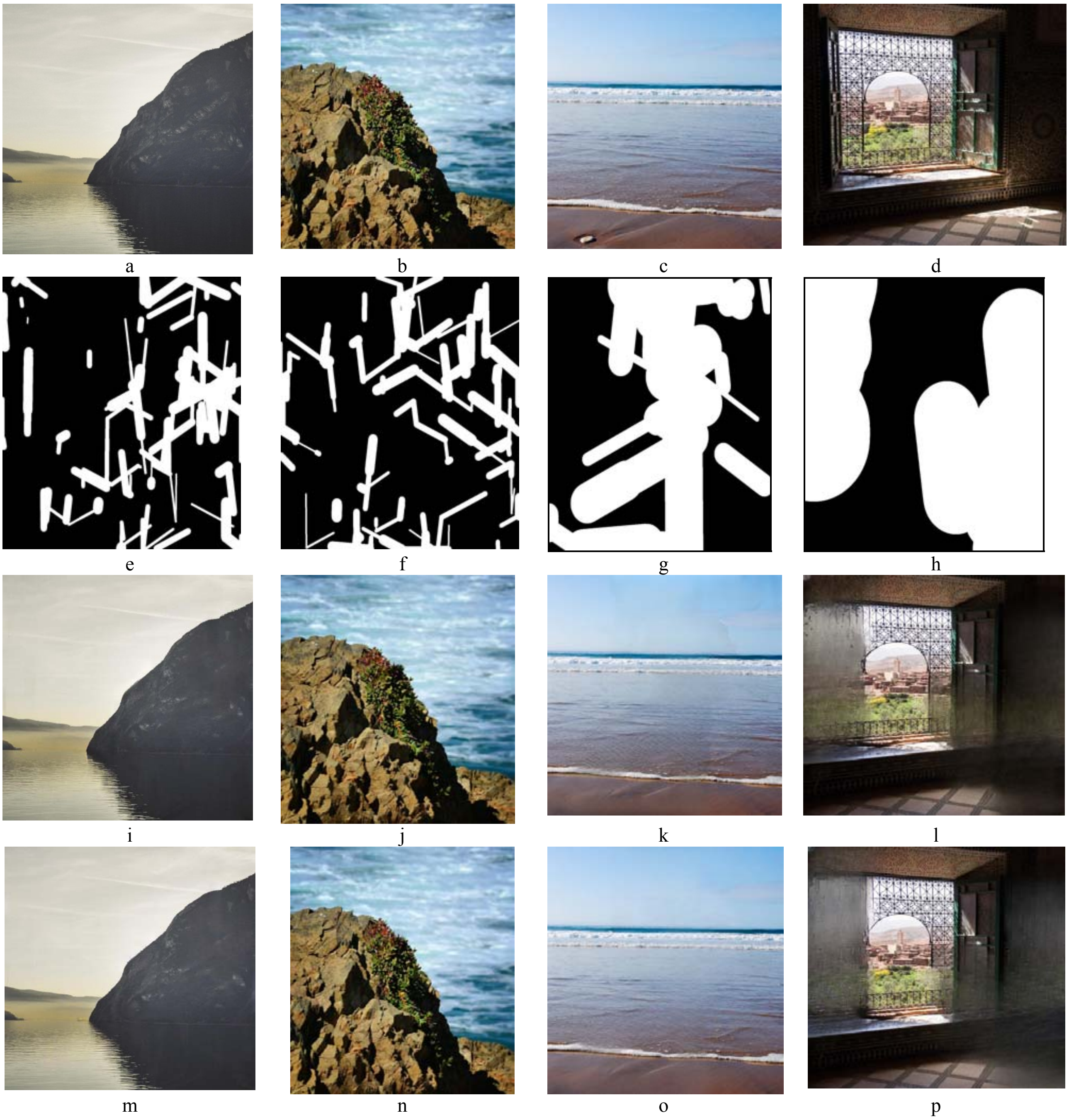


Figure 3 – The results of image inpainting with LaMa network: a, b, c, d – initial image; e, f – narrow mask; g – medium mask; h – large mask; i, j, k, l – images inpainted with LaMa-Fourier; m, n, o, p – images inpainted with LaMa-Wavelet

Prospect for further research is a reducing the prediction error. In this way, it is possible to apply the regression on several input parameters. To our opinion, the proposed method can be also applied to other CNNs trained to inpaint real scene images to decide about the advisability of the time and resource consumption needed for image inpainting.

ACKNOWLEDGEMENTS

The work was supported by the state budget research project of the Odessa Polytechnic National University "Modeling of self-organizing systems and controlled dynamic systems" (state registration number 0119U003520).

REFERENCES

1. Xiang H., Zou Q., Nawaz M. A. et al. Deep learning for image inpainting: a survey, *Pattern Recognition*, 2023, Vol. 134, Article 109046. DOI: 10.1016/j.patcog.2022.109046.
2. Xu Z., Zhang X., Chen W. et al. A review of image inpainting method based on deep learning, *Appl. Sci.*, 2023, Vol. 13, Article 11189. DOI: 10.3390/app132011189.

DOI 10.15588/1607-3274-2025-2-5

3. Ho J., Jain A., Abbeel P. Denoising diffusion probabilistic models, *Advances in Neural Information Processing Systems,* 2020, Vol. 33, pp. 6840–6851.
4. Lugmayr A., Danelljan M., Romero A. et al. Repaint: Inpainting using denoising diffusion probabilistic models, *Computer Vision and Pattern Recognition: IEEE/CVF Conference, New Orleans, LA, USA, 19–20 June 2022 : proceedings.* IEEE 2022, pp. 11451–11461. DOI: 10.1109/CVPR52688.2022.0111.
5. Yu J., Yang J., Shen X., Lu X., Huang T. S. Generative image inpainting with contextual attention. *Computer Vision and Pattern Recognition Workshops: IEEE/CVF Conference, CVPRW, Salt Lake City, UT, USA, 18–22 June, 2018 : proceedings.* IEEE, 2018, pp. 5505–5514. DOI: 10.1109/CVPRW.2018.00577.
6. Mohite T. A., Phadke G. S. Image inpainting with contextual attention and partial convolution, *Artificial Intelligence and Signal Processing: 2020 International Conference, AISP, Amaravati, India, 10–12 January 2020 : proceedings*. IEEE, 2020, pp. 1–6. DOI: 10.1109/AISP48273.2020.9073008.
7. Guo Q., Li X., Juefei-Xu F. et al. JPGnet: Joint predictive filtering and generative network for image inpainting, *Multimedia: 29th ACM International Conference, Chengdu, China, 20–24 October 2021 : proceedings.* ACM, 2021, pp. 386–394. DOI: 10.1145/3474085.3475170.
8. Suvorov R., Logacheva E., Mashikhin A. et al. Resolution-robust large mask inpainting with Fourier convolutions, *Applications of Computer Vision: IEEE Workshop/Winter Conference, WACV, Waikoloa, Hawaii, 4–8 January, 2022 : proceedings*. IEEE, 2022, pp. 2149–2159. DOI: 10.1109/WACV51458.2022.00323
9. Kolodochka D. O., Polyakova M. V. LaMa-Wavelet: image inpainting with high quality of fine details and object edges, *Radio Electronics, Computer Science, Control,* 2024, № 1, pp. 208–220. DOI: 10.15588/1607-3274-2024-1-19.
10. Jain S., Shivam V., Bidargaddi A. P., Malipatil S., Patil K. Image inpainting using YOLOv8 and LaMa model, *Emerging Technology: 5th International Conference (INCET), Belgaum, India, 24-26 May 2024 : proceedings.* IEEE, 2024, pp. 1–7. DOI: 10.1109/INCET61516.2024.10593536.
11. Kolodochka D., Polyakova M., Nesteriuk O., Makarichev V. LaMa network architecture search for image inpainting, *Information Control Systems & Technologies: 12th International Conference, ICST, Odesa, Ukraine, 23–25 September, 2024 : proceedings*. CEUR-WS, 2024, Vol. 3799, pp. 365 – 376.
12. Places365 Scene Recognition Demo [Electronic resource]. Access mode: http://places2.csail.mit.edu/
13. Mellor J., Turner J., Storkey A. J., Crowley E. J. Neural architecture search without training, *Machine Learning: 38th International Conference, PMLR, Virtual, 18–24 July 2021: proceedings*. IEEE, 2021, Vol. 139, pp. 7588–7598. DOI: 10.48550/arXiv.2006.04647.
14. Rubel O., Abramov S., Lukin V. et al. Is texture denoising efficiency predictable, *International Journal on Pattern Recognition and Artificial Intelligence,* 2018, Vol. 32, Article 1860005. DOI: 10.1142/S0218001418600054.
15. Rubel O. S., Lukin V. V., de Medeiros F. S. Prediction of despeckling efficiency of DCT-based filters applied to SAR images, *Distributed Computing in Sensor Systems: 2015 International Conference, (DCOSS), Fortaleza, Brazil, 10–12 June, 2015 : proceedings*. IEEE, 2015, pp. 159–168. DOI: 10.1109/DCOSS.2015.16.
16. Abramov S., Abramova V., Lukin V., Egiazarian K. Prediction of signal denoising efficiency for DCT-based filter, *Telecommunications and Radio Engineering*, 2019, Vol. 78, № 13, pp. 1129–114. DOI: 10.1615/TelecomRadEng.v78.i13.10.
17. Zalasiński M., Cader A., Patora-Wysocka Z., Xiao M. Evaluating neural network models for predicting dynamic signature signals, *Journal of Artificial Intelligence and Soft Computing Research,* 2024, Vol. 14, № 4, pp. 361–372. DOI: 10.2478/jaiscr-2024-0019.
18. Cao L., Yang T., Wang Y., Yan B., Guo Y. Generator pyramid for high-resolution image inpainting, *Complex & Intelligent Systems,* 2023, Vol. 9, Article 7553. DOI: 10.1007/s40747-023-01080-w.
19. Yamashita Y., Shimosato K., Ukita N. Boundary-aware image inpainting with multiple auxiliary cues, *Computer Vision and Pattern Recognition: IEEE/CVF Workshop/Conference, New Orleans, LA, USA, 19–20 June, 2022 : proceedings*. IEEE, 2022, pp. 618–628. DOI: 10.1109/CVPRW56347.2022.00077.
20. Nazeri K., Ng E., Joseph T., Qureshi F., Ebrahimi M. EdgeConnect: structure guided image inpainting using edge prediction, *Computer Vision Workshop: __IEEE/CVF International Conference, ICCVW, Seoul, Korea (South), 27–28 October, 2019 : proceedings.* IEEE, 2019, pp. 2462–2468. DOI: 10.1109/ICCVW.2019.00408
21. Liao L., Xiao J., Wang Z., Lin C.-W., Satoh S. Guidance and evaluation: semantic-aware image inpainting for mixed scenes, *Computer Vision: 16th European Conference, ECCV, Glasgow, UK, 23–28 August 2020 : proceedings*. IEEE, 2020, pp. 683–700. DOI: 10.1007/978-3-030-58583-9_41.
22. Liu G., Reda F. A., Shih K. J. et al. Image inpainting for irregular holes using partial convolutions, *Computer Vision: European Conference, ECCV, Munich, Germany, 8–14 September, 2018 : proceedings.* IEEE, 2018, pp. 85–100. DOI: 10.1007/978-3-030-01252-6_6.
23. Yu J., Lin Z., Yang J. et al. Free-form image inpainting with gated convolution, *Computer Vision: IEEE/CVF International Conference, ICCV, Seoul, Korea (South), 27 October – 2 November, 2019 : proceedings*. IEEE, 2019, pp. 4471–4480. DOI: 10.1109/ICCV.2019.00457.
24. Xie C., Liu S., Li C. et al. Image inpainting with learnable bidirectional attention maps, *Computer Vision: IEEE/CVF International Conference, ICCV, Seoul, Korea (South), 27 October – 2 November 2019 : procedings.* IEEE, 2019, pp. 8857–8866. DOI: 10.1109/ICCV.2019.00895.
25. Wang X., Girshick R., Gupta A., He K. Non-local neural networks, *Computer Vision and Pattern Recognition: 2018 IEEE/CVF Conference, Salt Lake City, UT, USA, 18–22 June 2018 : proceedings*. IEEE, 2018, pp. 7794–7803. DOI: 10.1109/CVPR.2018.00813.
26. Sara U., Akter M., Uddin M. S. Image quality assessment through FSIM, SSIM, MSE and PSNR – a comparative study, *Journal of Computer and Communications,* 2019, Vol. 7, № 3, pp. 8–18. DOI: 10.4236/jcc.2019.73002.
27. Gonzalez R. C., Woods R. E. Digital Image Processing, NY: Pearson, 4th Edition, 2017, 1192 p.

Received 05.02.2025.
Accepted 22.04.2025.

DOI 10.15588/1607-3274-2025-2-5

УДК 004.93

ПРОГНОЗУВАННЯ ЯКОСТІ ВІДНОВЛЕННЯ ЗОБРАЖЕНЬ ІЗ ЗАСТОСУВАННЯМ ТЕСТУРНИХ ДЕСКРИПТОРІВ

Колодочка Д. О. – студент Інституту комп'ютерних систем Національного університету «Одеська політехніка», Одеса, Україна.

Полякова М. В. – д-р техн. наук, доцент, професор кафедри прикладної математики та інформаційних технологій Національного університету «Одеська політехніка», Одеса, Україна.

Рогачко В. В. – студент Інституту комп'ютерних систем Національного університету «Одеська політехніка», Одеса, Україна.

АНОТАЦІЯ

Актуальність. Проблема заповнення відсутніх областей зображення реалістичним контентом часто виникає при обробці реальних сцен у комп'ютерному зорі та комп'ютерній графіці. Щоб відновити відсутні області на зображенні, застосовуються різні підходи, такі як дифузійні моделі, механізм самоуважності, генеративні змагальні мережі. Для відновлення зображень реальних сцен використовуються згорткові нейронні мережі. Із застосуванням цих мереж останнім часом досягнуто значних успіхів у відновленні зображень. Але отримані відновлені зображення не завжди високої якості.

Мета роботи полягає у зменшенні витрат часу в системах комп'ютерної графіки та комп'ютерного зору шляхом прогнозування якості відновлення зображень згортковими нейронними мережами.

Метод. Прогноз точності відновлення зображення здійснено шляхом аналізу статистики зображення без виконання самої реконструкції і, отже, без витрачання зайвого часу та комп'ютерних ресурсів на відновлення зображення. Ми використали пікове відношення сигнал/шум і показник індексу структурної подібності для оцінки якості відновлення зображення.

Результати. Показано, що передбачення ефективне для широкого діапазону розмірів масок і зображень реальних сцен з бази даних Places2. У якості прикладу було зосереджено на окремих випадках версій мережі LaMa, хоча запропонований метод також можна узагальнити на інші згорткові нейронні мережі.

Висновки. Отримані результати показують, що прогноз якості відновлення зображень може бути виконаний із задовільною точністю, якщо використовувати залежності SSIM або PSNR від показника однорідності текстури зображень. Слід зазначити, що структурна подібність початкового та відновленого зображень краще передбачувана, ніж помилка між відповідними пікселями цих зображень. Щоб зменшити помилку прогнозування можна застосувати регресію за декількома вхідними змінними.

КЛЮЧОВІ СЛОВА: відновлення зображення, прогнозування точності, мережа LaMa, дескриптор текстури, матриця суміжності.

ЛІТЕРАТУРА

1. Deep learning for image inpainting: a survey / [H. Xiang, Q. Zou, M. A. Nawaz et al.] // Pattern Recognition. – 2023. – Vol. 134. – Article 109046. DOI: 10.1016/j.patcog.2022.109046.
2. A review of image inpainting method based on deep learning / [Z. Xu, X. Zhang, W. Chen et al.] // Applied Sciences. – 2023. – Vol. 13. – Article 11189. DOI: 10.3390/app132011189.
3. Ho J. Denoising diffusion probabilistic models / J. Ho, A. Jain, P. Abbeel // Advances in Neural Information Processing Systems. – 2020. – Vol. 33. – P. 6840–6851.
4. Repaint: Inpainting using denoising diffusion probabilistic models / [A. Lugmayr, M. Danelljan, A. Romero et al.] // Computer Vision and Pattern Recognition: IEEE/CVF Conference, New Orleans, LA, USA, 19–20 June 2022 : proceedings. – IEEE 2022. – P. 11451–11461. DOI: 10.1109/CVPR52688.2022.0111.
5. Generative image inpainting with contextual attention / [J. Yu, J. Yang, X. Shen et al.] // Computer Vision and Pattern Recognition Workshops: IEEE/CVF Conference, CVPRW, Salt Lake City, UT, USA, 18–22 June, 2018 : proceedings. – IEEE, 2018. – P. 5505–5514. DOI: 10.1109/CVPRW.2018.00577.
6. Mohite T. A. Image inpainting with contextual attention and partial convolution / T. A. Mohite, G. S. Phadke // Artificial Intelligence and Signal Processing: 2020 International Conference, AISP, Amaravati, India, 10–12 January 2020 : proceedings. – IEEE, 2020. – P. 1–6. DOI: 10.1109/AISP48273.2020.9073008.
7. JPGnet: Joint predictive filtering and generative network for image inpainting / [Q. Guo, X. Li, F. Juefei-Xu et al.] // Multimedia: 29th ACM International Conference, Chengdu, China, 20–24 October 2021 : proceedings. – ACM, 2021. – P. 386–394. DOI: 10.1145/3474085.3475170.
8. Resolution-robust large mask inpainting with Fourier convolutions / [Suvorov R., Logacheva E., Mashikhin A. et al.] // Applications of Computer Vision: IEEE Workshop/Winter Conference, WACV, Waikoloa, Hawaii, 4 – 8 January, 2022 : proceedings. – IEEE, 2022. – P. 2149–2159. DOI: 10.1109/WACV51458.2022.00323.
9. Kolodochka D. O. LaMa-Wavelet: image inpainting with high quality of fine details and object edges / D. O. Kolodochka, M. V. Polyakova // Radio Electronics, Computer Science, Control. – 2024. – № 1. – P. 208–220. DOI: 10.15588/1607-3274-2024-1-19.
10. Image inpainting using YOLOv8 and LaMa model / [S. Jain, V. Shivam, A. P. Bidargaddi et al.] // Emerging Technology: 5th International Conference (INCET), Belgaum, India, 24–26 May 2024 : proceedings. – IEEE, 2024. – P. 1–7. DOI: 10.1109/INCET61516.2024.10593536.
11. LaMa network architecture search for image inpainting / [D. Kolodochka, M. Polyakova, O. Nesteriuk, V. Makarichev] // Information Control Systems & Technologies: 12th International Conference Information, ICST, Odesa, Ukraine, 23–25 September, 2024 : proceedings. – CEUR-WS, 2024. – Vol. 3799. – P. 365 – 376.
12. Places365 Scene Recognition Demo [Electronic resource]. Access mode: http://places2.csail.mit.edu/

DOI 10.15588/1607-3274-2025-2-5

13. Neural architecture search without training / [J. Mellor, J. Turner, A. J. Storkey, E. J. Crowley] // Machine Learning: 38th International Conference, PMLR, Virtual, 18–24 July 2021: proceedings. – IEEE, 2021. – Vol. 139. – P. 7588–7598. DOI: 10.48550/arXiv.2006.04647.
14. Is texture denoising efficiency predictable / [O. Rubel, S. Abramov, V. Lukin et al.] // International Journal on Pattern Recognition and Artificial Intelligence. – 2018. – Vol. 32. – Article 1860005. DOI: 10.1142/S0218001418600054.
15. Rubel O. S. Prediction of despeckling efficiency of DCT-based filters applied to SAR images / O. S. Rubel, V. V. Lukin, F. S. de Medeiros // Distributed Computing in Sensor Systems: 2015 International Conference (DCOSS), Fortaleza, Brazil, 10 – 12 June, 2015 : proceedings. – IEEE, 2015. – P. 159–168. DOI: 10.1109/DCOSS.2015.16.
16. Prediction of signal denoising efficiency for DCT-based filter / [S. Abramov, V. Abramova, V. Lukin, K. Egiazarian] // Telecommunications and Radio Engineering. – 2019. – Vol. 78, № 13. – P. 1129–114. DOI: 10.1615/TelecomRadEng.v78.i13.10.
17. Evaluating neural network models for predicting dynamic signature signals / [M. Zalasiński, A. Cader, Z. Patora-Wysocka, M. Xiao] // Journal of Artificial Intelligence and Soft Computing Research. – 2024. – Vol. 14, № 4. – P. 361–372. DOI: 10.2478/jaiscr-2024-0019.
18. Generator pyramid for high-resolution image inpainting / [L. Cao, T. Yang, Y. Wang, B. Yan, Y. Guo] // Complex & Intelligent Systems. – 2023. – Vol. 9. – Article 7553. DOI: 10.1007/s40747-023-01080-w.
19. Yamashita Y. Boundary-aware image inpainting with multiple auxiliary cues / [Y. Yamashita, K. Shimosato, N. Ukita] // Computer Vision and Pattern Recognition: IEEE/CVF Workshop/Conference, New Orleans, LA, USA, 19 – 20 June, 2022 : proceedings. – IEEE, 2022. – P. 618–628. DOI: 10.1109/CVPRW56347.2022.00077.
20. EdgeConnect: structure guided image inpainting using edge prediction / [K. Nazeri, E. Ng, T. Joseph et al.]// Computer Vision Workshop: _IEEE/CVF International Conference, ICCVW, Seoul, Korea (South), 27–28 October, 2019 : proceedings. – IEEE, 2019. – P. 2462–2468. DOI: 10.1109/ICCVW.2019.00408.
21. Guidance and evaluation: semantic-aware image inpainting for mixed scenes / [L. Liao, J. Xiao, Z. Wang et al.] // Computer Vision: 16th European Conference, ECCV, Glasgow, UK, 23–28 August 2020 : proceedings. – IEEE, 2020. – P. 683–700. DOI: 10.1007/978-3-030-58583-9_41.
22. Image inpainting for irregular holes using partial convolutions / [G. Liu, F. A. Reda, K. J. Shih et al.] // Computer Vision: European Conference, ECCV, Munich, Germany, 8 – 14 September, 2018 : proceedings. – IEEE, 2018. – P. 85–100. DOI: 10.1007/978-3-030-01252-6_6.
23. Free-form image inpainting with gated convolution / [J. Yu, Z. Lin, J. Yang et al.] // Computer Vision: IEEE/CVF International Conference, ICCV, Seoul, Korea (South), 27 October – 2 November, 2019 : proceedings. – IEEE, 2019. – P. 4471–4480. DOI: 10.1109/ICCV.2019.00457.
24. Image inpainting with learnable bidirectional attention maps / [C. Xie, S. Liu, C. Li et al.] // Computer Vision: IEEE/CVF International Conference, ICCV, Seoul, Korea (South), 27 October – 2 November 2019 : procedings. – IEEE, 2019. – P. 8857–8866. DOI: 10.1109/ICCV.2019.00895.
25. Non-local neural networks / [X. Wang, R. Girshick, A. Gupta, K. He] // Computer Vision and Pattern Recognition: 2018 IEEE/CVF Conference, Salt Lake City, UT, USA, 18 – 22 June 2018 : proceedings. – IEEE, 2018. – P. 7794–7803. DOI: 10.1109/CVPR.2018.00813.
26. Sara U. Image quality assessment through FSIM, SSIM, MSE and PSNR — a comparative study / U. Sara, M. Akter, M. S. Uddin // Journal of Computer and Communications. – 2019. – Vol. 7, № 3. – P. 8–18. DOI: 10.4236/jcc.2019.73002.
27. Gonzalez R. C. Digital Image Processing / R. C. Gonzalez, R. E. Woods. – NY: Pearson, 4th Edition, 2017. – 1192 p.

DOI 10.15588/1607-3274-2025-2-5