Inpainting of Irregular Holes in a Manuscript using UNet and Partial Convolution

Amreen Kaur, Ankit Raj, N. Jayanthi, S.Indu
Department of Electronics and Communication Engineering
Delhi Technological University (Formerly Delhi College of Engineering), Delhi 110042, India
amreenrocks.98@gmail.com (Amreen Kaur), rajankit606@gmail.com (Ankit Raj),
njayanthi@dce.ac.in (N. Jayanthi), s.indu@dce.ac.in(S.Indu)

Abstract - **Image inpainting of irregular holes in the old handwritten manuscripts is a challenging task. It requires finding the missing part in the manuscript and then filling it with an appropriate pixel. It needs to fill it in such a way that words or letters in that manuscript can be recovered. There are many degraded manuscripts available in museums and libraries all across the globe. This degradation can be due to human mishandling or certain environmental factors (fire, heat, water, microorganisms, pests, and other vermin). Even if it is tired very hard to protect our written heritage from all this, since it is written on natural materials (wood, paper), it is prone to inherent vice. A novel method is presented to inpaint the irregular holes in manuscripts with appropriate pixels. Deep learning, which tends to solve many complex problems like image classification, object detection, image segmentation, helps in achieving image inpainting as well. UNet is proposed to generate the mask of an irregular hole part in a manuscript image and partial convolution, which takes two input masks and the image for inpainting.**

Keywords—Image inpainting, UNet, deep learning, partial convolution, manuscript.

I. INTRODUCTION

Manuscripts are the witness of the past. From the moment human beings figured out the need to keep a note of things, manuscripts came into existence. Manuscripts that have existed in antiquity for centuries have considerable significance. Famous personalities in the past used to write manuscripts to keep a record of their thoughts, learnings, and incidents that happened around them. These literary sources help historians and archaeologists to draw a clear picture of the ancient times. Basically, the timeline leading us till very recent times can be framed by them. So, very naturally these hold immense value to people.

Such manuscripts are links in the advancement of people of an area and of humanity, whether for the purpose of technology or for the transmission of information. And though the original had deteriorated, most of these manuscripts were physically copied by scribes, and in effect, the copied manuscript was recopied before reaching us. If people back then were trying to put in such huge efforts to keep the texts alive and maintain their longevity, their hardwork should definitely be not wasted and also preserve them for our own goods and benefits. But being under the weather for years, these documents serve no use unless treated properly. Prevention and digital archiving of historical documents are increasing [23]. These documents suffer from many forms of distortion, degradation, color loss, physical degradation, insect bites, dust, disasters, low-quality paper, etc. Archived images required enhancement and restoration [14] using various techniques to obtain its original look. Restoration helps to extract text and many features from images and helps us preserve manuscripts. Fig.1 depicts the sample of a damaged Indian Manuscript which can be corrected with our model.

Fig 1: A Degraded Manuscript with irregular holes.

Image inpainting is a process of filling up degradation in images or reconstructing them. Earlier it was done by humans but now Deep Learning techniques have replaced them. It produces images such that the existence of damage wouldn't be understood. In these methods, the input is required as a mask showing the regions that should be inpainted. So, old manuscripts require filling up irregular holes in them, in order to form documents from which historians can infer useful data. Not a lot of Deep Learning techniques exist which tackle the problem of image restoration. One approach is using stacked up partial convolution functions while the mask is being updated [1]. Restoration of old manuscripts through inpainting methods will result in a more accurate restoration of manuscripts.

This paper tries to achieve the segmentation of irregular holes using the U-Net [8] approach. Olaf Ronneberger [8] used U Net for the segmentation of neuronal structure during the ISBI challenge. He shows that the network can be trained from a few images by strong use of data augmentation. The architecture is fast and outperforms which shows precise localization. It can also be defined as a neural network designed as an encoder-decoder which solves end to end segmentation tasks. There exists an algorithm related to supervise U-Net. This method is utilized for automatic segmentation in the field of medicine [3].

Apparently UNet was optimized to have importance in the life sciences genre. UNet also finds use in Iris Segmentation. Iris Segmentation has its need in medical sectors as well as in the biometric division. The plus point of using UNet here is its ability to comprehend accurate models from a rather small dataset [4].

Apart from UNet, Partial Convolution [1] covers a significant part of this paper which achieved an impressive result in image inpainting. Dynamic Pacemaker Artifact Reduction [5]. Another problem solved by Partial Convolution (in place of traditional convolution methods) is the restoration of digital Dunhuang murals. It forms a partial convolution layer with the help of the partial convolution operation and masks update function [6]. Image to image convolution neural networks [24] form the basis of Metal Artifact Reduction (MAR) techniques. This problem is seen in painting as a direct relation to a sonogram correction issue and hence partial convolution with UNet architecture is allowed to be used in a part of the pipeline [7].

The major task includes collecting the damaged dataset from all the possible resources and having wide variations, training the model using deep learning, and then testing it. Combining the UNet and the Partial Convulsion architectures, a new solution to the restoration of irregular holes in degraded documents was devised. This method is seemingly new and it has not been explored before this. A mask of the irregular hole will be primarily created in the degraded document and then inpaint them using Partial convolution.

The organization of this paper is as follows. In the section that follows a literature review is provided. In the next section, detailed information related to the dataset is presented to deeply understand the problems that will be dealt to correctly identify the damages done and inprint the manuscript. A detailed description of the architecture employed to design the algorithm is then presented. The process shall continue until the most accurate results possible are obtained for the model. The results and conclusion are then mentioned at the end of the paper.

II. LITERATURE REVIEW

Over the years studies and experimentation methods to extract text from a variety of handwritten, degraded, or historical documents have surfaced. Recently it has been suggested that low contrast texts from the degraded documents can be figured out from the intersection operation of Laplacian and Sobel edge images. [15] All pixels around the edges are enhanced by the Laplacian Operation. Noise pixels are produced, as a result of background variation. On the other hand, the Sobel operator carries out the enhancement of the high contrast pixel without generating noise pixels. The intersection of the operations, result in sieving out the pixels which are relevant for text line segmentation. Further, the width of the edge pixels is reduced by the use of the skeleton. A clustering based on the nearest neighbor criterion is used to segment the text lines. The cluster with lesser branches is the text cluster. This method was devised for transcription of documents from the Indus Valley Civilization. Document image binarization is a useful approach, to separate the background from the foreground by attaching 0 and 1 values to ink and non-ink elements. The most recently developed method to achieve the same is BiNet [16]. This method has been inspired by the U-Net architecture. It uses deep-encoder decoder networks. Here the function is a mapping from the input image to the output image: $f(x) = x + \Delta$, where x is the DSS (Dead Sea Scrolls image collection), image input, and $\Delta$is the noise (information other than the original content). The classification error of the loss function is then minimized and a regression line is plotted for the function to give the prediction.

Another one is the OCR less approach [17]. It proceeds with a variety of image processing techniques followed by de-skewing of images. Image Quality metrics are used for the evaluation of the obtained original text document from the degraded input. The process of deskewing detects and extracts the SURF feature. The geometric transformation between skewed and deskewed images is estimated. The image is converted to Grayscale from the RGB, using the format YCbCr for image restoring and then normalization is applied. In this technique, luminance comparison is done and the values are copied from the RGB image to the Grayscale image.

Restoration is another important aspect of the digitization of degraded documents. Initially global and local thresholding methods are applied. The Niblack method is used to calculate the local threshold of each pixel. $T(x, y) = m(x, y) + k * s(x, y)$, where m is the average of the local area, and s is the standard deviation, and k adjusts the boundary. Acute changes to the spatial averaging are done by the Gaussian Filtering Method [18]. The sigma of the Gaussian distribution gives the degree of noise reduction. Gabor transformation accounts for the preprocessing task, followed by the preprocessing done by the Gaussian Filtering Method. Finally, the edges are detected and the holes are filled. The grey level gradient at a pixel is estimated to make the background uniform and remove the stains.

Another approach makes use of the Markovian-Bayesian clustering [19]. The process is composed of three steps. The image is modeled by 4-nearest neighbors. In the first step, parameters of the Gaussian model are estimated by the iterative conditional estimation (ICE). Hence the result proves the over-segmentation behavior, by dividing the region into numerous regions. The next step in the algorithm is the regions merging based on the degree of homogeneity. Bhattacharya distance is used in this regard. Finally, binarization with k-means clustering is done which goes on to achieve text segmentation of the document.

Multispectral imaging [20] has also gained some prominence in the field of document restoration. This method is based on the concept of pattern recognition. The proposed model is an unsupervised restoration model. The iron-gall ink is chosen as it gives different results in infrared channels and the degradation continues to be visible. When the degradation and the text overlap the end-members extraction is used to isolate the text pixels. After removing these pixels from the mask, the inpainting is then applied to the refined mask thus formed. This method effectively preserves the quality of the document as well as the original view.

Decompose algorithm is another approach of looking into the same problem, where it is started by detecting the presence of a Bimodal Histogram [21]. If the answer is affirmative the global thresholding method (time complexity being O(n)) with the highest accuracy is applied to it. This method is vividly effective in case there isn't any non-bimodal histogram present. It then focuses on the 8-bit grayscale of the document. Local feature vectors are then used for thresholding a small area. These local regions are then classified and their edges are smoothened. Finally, thresholding is applied to each region. In the case of the absence of bimodality in the image, a recursive decomposition approach takes place using quadtrees. A thresholding method based on the weighted gradient is used once the entire image has been decomposed. The complexity of this one is $O(n^2)$.

Another approach is to divide or segregate a degraded image into the background, original text, and interfering text (which is then replaced by the average of the background pixel) [21]. A segmentation approach is applied recursively on decorrelated data as it is most unlikely to achieve the result by single clustering. On generated data, recursion is applied to Principal Component Analysis (PCA) and K-means algorithms. PCA is a dimensionality reduction algorithm based on eigenvectors. The k-means algorithm is used which divides the image pixels into two classes.

Mostly, the restoration problem is an unsupervised learning problem and newer algorithms are coming up. But to date, there isn't any structured algorithm to completely solve the degraded document issue. This has been primarily aimed in our paper.

III. METHODOLOGY

The aim is to inpaint the holes or damaged parts inside a document. A two-stage process is presented. The first stage does segmentation of irregular holes and creates a mask using U Net. The mask is then dilated [14]. To remove the boundary of irregular holes. The Second stage does Partial convolution which takes an image and a mask as input and produces a new reconstructed Image. The fig.2 depicts the architecture of our model.

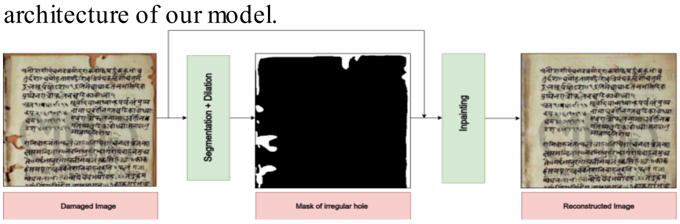

Fig.2: Architecture of our model

3.1 *Dataset*

The role of the dataset in this paper is massive. The primary source of the dataset was the internet where a few sites provide an open and usable database of manuscripts. Apart from the Internet, the ancient Indian Manuscripts are collected from museums, history textbooks, and from the Library of Delhi Technological University. It contains Sanskrit, Prakrit, Gurumukhi, Avadhi Languages along with the Medieval Paleographic Scale (MPS) [9] dataset.

The quality of the documents varied ranging from different handwritings to different page qualities all of which were taken into account during the preprocessing of the images of the manuscripts. The images are resized and cropped into 512 x 512. The dataset is categorized into the damaged and non-damaged dataset. The damaged dataset is used to create a mask and then train the UNet Network. The non-damaged dataset and random mask generator are used to train the partial convolution network. Data Augmentation technique is applied to the dataset to generate more document images, to help train the network with fewer images. Table-1 shows the data in tabular form.Fig.3 shows the sample images of the dataset.

Table 1: Description of dataset

Language	No.of Images	Non-Damaged	Damaged
Sanskrit	1200	896	304
Prakrit	1200	923	277
Punjabi	500	402	98
Avadhi	250	205	45
MPS dataset	2500	1986	514

a) Prakrit b) Sanskrit c) Avadhi d) Punjabi e) MPS

Fig.3: Sample images from the Non-Damaged dataset

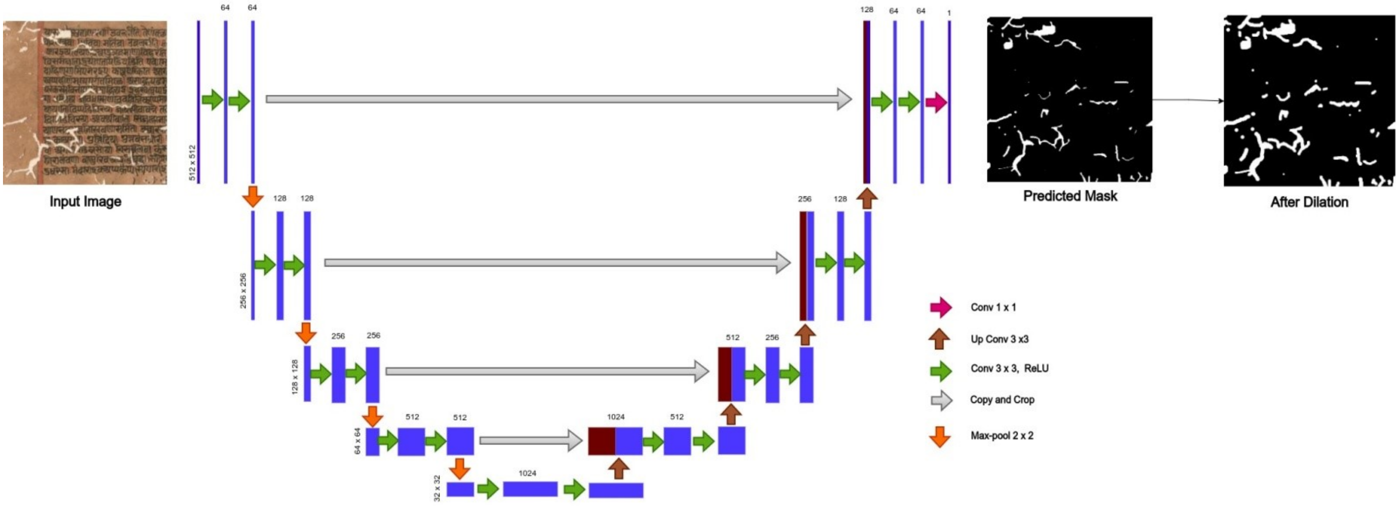

Fig.4: UNet Architecture of our model

3.2 *UNet*

The U-Net architecture [8] is a U-shaped Fully Convolutional Network (FCN) that results in the precise segmentation of the image. The improvement on FCN with the emergence of UNet is that UNet is a symmetric design and there is a skip connection between the downsampling paths. Moreover, the operator used in the upsampling path is concatenation and not sum. It is a symmetric network that is made up of downsampling and upsampling paths. It also contains a great number of feature maps to pass the required information via the upsampling route. UNet has its own advantage of not having dense layers. It makes it suitable for various sizes of image input. Massive data augmentation is also processed efficiently in the UNet.Fig.4 depicts the UNet Architecture. UNet architecture consists of three parts:

A. The contracting/downsampling path:

Recursive application of convolution creates the contracting path. This path contains 4 blocks. Each block has two 3x3 Convolution Layers + ReLU activation functions (with batch normalization), and a 2x2 Max Pooling layer. After the execution of every pooling functioning process, the value of the number of feature maps increases by double. The pooling layer ensures that the number of channels going to the input matrix is always fixed. There is an increase in feature information and a decrease in the spatial ones while contraction takes place. The intention here is to be familiar with the context of the image that is given as input. This enables a better segmentation. This information is then passed to an upsampling way.

B. Bottleneck:

The intention here is to be familiar with the context of the image that is given as input. This enables a better segmentation. This information is then passed to an upsampling way.

C. The expanding/upsampling path:

A combination of the feature and spatial information occurs in the expansive path. It happens by sequential execution of up-convolutions and concatenation operations. The upsampling part also contains 4 blocks. Each block has a Deconvolution layer with two 3x3 Convolution layers + ReLU activation function (with batch normalization).

As a result of this, the upsampling path is almost symmetric to the downsampling path and gives us a u-shaped architectural structure. The task of the U Net architecture is to perform the classification of each pixel of the input image and label it. The final output consists of pixel-wise softmax-max of the entire final feature map with a cross-entropy loss function.

The soft-max is expressed as:

$$p_k(x) = \frac{exp(a_k(x))}{\sum_{K'=1}^{K} \exp(a_{k'}(x))} \quad (1)$$

$a_k(x) = activation\ in\ feature\ channel\ k\ at\ the\ pixel\ position$

$K = Number\ of\ classes (in\ our\ model\ K = 2)$

$p_x(x) = approximated\ maximum\ function$

The cross-entropy at each position $p_{i(x)}(x)$ is written as:

$$E = \sum_{x \propto \Omega} w(x) \log(p_{l(x)}(x)) \quad (2)$$

$\Omega \rightarrow \{0,1\}\ true\ label\ of\ each\ pixel$

Olaf Ronneberger [8] highly uses data augmentation with few mages that help to train the model with few datasets. There is also a few irregular images. So, this research work has used the UNet network and highly used data augmentation to train our model.

Fig. 5: i)The first row shows damaged images ii) The second row shows the mask of irregular holes iii) The last row shows the reconstructed image.

3.3 Partial Convolution

This model consists of stacked partial convolution operations and updating masks to achieve image inpainting. A partial convolution along with a mask update in place of traditional convolution layers gives a much better result in image inpainting involving irregular holes. The partial convolution operation is expressed as:

$$x' = \begin{cases} W^T(X \odot M)\frac{1}{sum(M)} + b\,, & if\ sum(M > 0) \\ 0, & otherwise \end{cases} \quad (3)$$

$W =$ *convolution filter weights for the convolution filter*
$b =$ *corresponding bias.*
$X =$ *the feature values (pixels values) for the current convolution (sliding) window.*
$M =$ *is the corresponding binary mask.*
$\odot =$ *element − wise multiplication.*

The output values depend only on the unmasked inputs. The mask is updated after each partial convolution operation. The mask M is updated as:

$$m' = \begin{cases} 1, & if\ sum(M) > 0 \\ 0, & otherwise \end{cases} \quad (4)$$

The architecture is similar to U Net, restoring all convolutional layers by the partial convolutional layer and nearest neighbor up-sample during decoding stages. The proposed work has used the same architecture as proposed by Guilin Liu [1]. According to [1] loss function is express as:

$$L_{total} = L_{valid} + 6L_{hole} + 0.05L_{perceptual} + 120(L_{style_{out}} + L_{style_{comp}} + 0.1L_{tv}) \quad (5)$$

L_{total} = *Total loss*
L_{valid}, L_{hole} = *Per pixel loss of the image*
$L_{style_{out}}, L_{style_{comp}}$ = *Style loss*
L_{tv} = *Total variantion loss*

IV. EXPERIMENTS AND RESULTS

The dataset is further divided into three-parts- training, validation, and testing. 60 % of data was used for training, 20 % for validation, and 20 % for testing. All the images are of size 512 x 512.

4.1 Masking

As shown in fig 1 and fig 5 the color of irregular holes is mostly white. The colored damaged image is converted into grayscale. The Non-hole part pixels in an image are low and the holes in the images have high pixels. So a threshold point is randomly selected and this helps in the binarization of the image. The different images have different threshold points. Some holes are not white in color but their pixel value is higher than the non-damaged parts of the image.

4.2 Training Process

This process has used damaged images and their manually annotated masks to train the U Net model using Adam and binary cross-entropy with up to 200 epochs. For training the Partial Convolution, random masks of size 512 x 512 are generated using OpenCV consisting of rectangles, circles, lines. Further this process will use these masks and the non-damaged dataset to then train the partial convolution network. This research work has started with the initial weight [11] and Adam [12] Optimization. Holes create a problem for Batch Normalization due to the mean and variance computed for hole pixels. To solve the mean and variance problem, the model is trained in two steps, which also helps to gain faster convergence. The model is trained in two steps first with initial training with enabled Batch Normalization with the learning rate of 0.0002 with up to 120 epochs and then fine-tuning with disabled Batch Normalization in the encoder with the learning rate of 0.00005 with up to 80 epochs. The batch size is 5 and 2500 steps per epoch in both cases. Fig.6 shows the corresponding results with the predicted output after masking on non-damaged dataset images with random lines, circles, and rectangles.

Fig.6: i) Masked Image ii) predicted image iii) Original image

4.3 Performance on the degraded dataset

After Training, the proposed model is tested on the damaged dataset to inpaint its irregular holes. The proposed model is tested for every angle of damaged images and for small irregular holes to larger irregular holes. Fig.5 shows the damaged images, their masks, and the corresponding predicted outputs.

4.4 Qualitative Analysis

Different square masks of sizes-16 x 16, 32 x 32, 64 x 64, 128 x 128 and straight-line masks of pixel 4 x 512, 8 x 512, 16 x 512, 32 x 512 on a 512 x 512 sized image are created. These masks are then added on non-damaged images and then image prediction is done in our model. Two Quantitative approaches are used to compare the predicted image and original image using Peak Signal to Noise Ratio(PSNR) which shows the ratio of the actual signal and the noise in the image, and Structural Symmetry Index(SSIM) which tells the similarity index of two images as shown in Table 2 and 3.

Table 2: PSNR AND SSIM values for each of the square masks

	16 x 16	32 x 32	64 x 64	128 x 128
PSNR	38.17	37.46	32.528	26.05
SSIM	0.972	0.971	0.965	0.935

Table 3: PSNR AND SSIM values for each of the straight-line masks

	4 x 512	8 x 512	16 x 512	32 x 512
PSNR	36.83	35.85	34.67	29.78
SSIM	0.971	0.966	0.964	0.94

Fig.7: i) Damaged image ii) Damaged mask iii) Reconstructed image

V. DISCUSSION

This paper proposes the U Net and partial convolution model to create an automatic mask and then automatic mask update to inpaint it. This model can easily handle small irregular holes, small lines, and recover the text with utmost perfection. But, it somehow does not show the best results and fails to recover text and inpaint the larger holes. Fig 7 shows some error due to dilation and not proper segmentation of irregularly large holes.

VI. CONCLUSION

This paper is a novel initiative to help restore degraded old Indian Manuscripts using UNet and partial convolution techniques. The specific work is done in this paper ie. of inpainting holes in a manuscript using Deep learning algorithms show promise with regard to computational performance. Concerning the accuracy, the network did perform well. It was efficiently able to inpaint irregular holes and tears of small sizes. However, it did face some difficulty in tackling larger holes and damages due to the dilation of the images. On using different datasets the architecture applied still proves to be very accurate every time.

VII. FUTURE WORK

IN future, it will include a wider dataset and irregular mask dataset [1] for our model to make it more robust. Also, the problem of error due to dilation that occurred in manuscripts along with larger holes will be catered to, in order to make the digitization of manuscripts easier and efficient. Nevertheless, it will also use other inpainting and segmentation algorithms to make the model to remain more accurate.

Acknowledgement

We would like to thank Delhi Technological University and our friends and family who helped us in finalizing the paper. We are sincerely thankful to www.indianmanuscripts.com from where we have collected most of our dataset.

References

[1] Guilin Liu, Fitsum A. Reda, Kevin J. Shih, Ting-Chun Wang, Andrew Tao, Bryan Catanzaro, Image Inpainting for Irregular Holes Using Partial Convolutions. The European Conference on Computer Vision (ECCV), 2018, pp. 85-100.

[2] Falk, T., Mai, D., Bensch, R., Çiçek, Ö., Abdulkadir, A., Marrakchi, Y., Böhm, A., Deubner, J., Jäckel, Z., Seiwald, K., Dovzhenko, A., Tietz, O., Bosco, C.D., Walsh, S., Saltukoglu, D., Tay, T.L., Prinz, M., Palme, K., Simons, M., Diester, I., Brox, T., Ronneberger, O., 2018. U-Net: deep learning for cell counting, detection, and morphometry.

[3] Li, S., Tso, G.K., He, K., 2020. Bottleneck features supervised U-Net for pixel-wise liver and tumor segmentation. Expert Systems with Applications 145, 113131.

[4] Lozej, J., Meden, B., Struc, V., Peer, P., 2018. End-to-End Iris Segmentation Using U-Net. 2018 IEEE International Work-Conference on Bioinspired Intelligence (IWOBI).

[5] Elss T.L., N., Nickisch, H., Wissel, T., Morlock, M., Grass, M., 2020. Learning metal artifact reduction in cardiac CT images with moving pacemakers. Medical Image Analysis 61, 101655.

[6] Chen, M., Zhao, X., Xu, D., 2019. Image Inpainting for Digital Dunhuang Murals Using Partial Convolutions and Sliding Window Method. Journal of Physics: Conference Series 1302, 032040.

[7] Samoylenko, A., Golanov, A., Pimkin, A., Antipina, N., Ovechkina, A., Daechina, A., Belyaev, M., 2019. Multi-Domain CT Metal artifacts reduction using partial convolution-based inpainting.

[8] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In MICCAI, pages 234–241. Springer, 2015

[9] Sheng He, Petros Samara, Jan Burgers, Lambert Schomaker. Multiple-Label Guided Clustering Algorithm for Historical Document Dating and Localization IEEE Trans. on Image Processing, Vol. 25(11), Nov. 2016. http://ieeexplore.ieee.org/document/7551181/

[10] Jonathan Long, Evan Shelhamer, Trevor Darrell, UC Berkeley, Fully Convolutional Networks for Semantic Segmentation.

[11] He, K., Zhang, X., Ren, S., Sun, J.: Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: Proceedings of the IEEE international conference on computer vision. pp. 1026–1034 (2015)

[12] Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)

[13] M.Raid, al, Image Restoration Based on Morphological Operations, International Journal of Computer Science, Engineering and Information Technology (IJCSEIT), Vol. 4, No.3, June 2014.

[14] Yahya, S. R., Abdullah, S. N. H. S., Omar, K., Zakaria, M. S., Liong, C. Y., "Review on image enhancement methods of the old manuscript with damaged background," in International Conference on Electrical Engineering and Informatics, pp. 62-67, 2009.

[15] Kavitha, A., Shivakumara, P., Kumar, G., Lu, T., 2016. Text segmentation in degraded historical document images. Egyptian Informatics Journal 17, 189–197.

[16] Dhali, M.A., Wit, J.W., Schomaker, L. Binet: Degraded-Manuscript Binarization in Diverse Document Textures and Layouts using Deep Encoder-Decoder Networks. [eprint] 2019.

[17] Hamsaveni L., Prakash N., Suresha., 2017. Degraded Document Analysis and Extraction of Original Text Document: An Approach without Optical Character Recognition. International Journal of Computer and Information Engineering 11, 110-114.

[18] Raj V., Arunkumar C., 2014. Content Restoration of Degraded Termite Bitten Document Images. International Journal of Computer and Information Engineering 4, 151-155.

[19] Hedjam R., Moghaddam R. F., Cheriet M., Text Extraction from degraded document images, 2010 2nd European Workshop on Visual Information Processing (EUVIP), Paris, 2010, 247-252.

[20] Hedjam, R., Cheriet, M., 2013. Historical document image restoration using a multispectral imaging system. Pattern Recognition 46, 2297–2312.

[21] Chen, Y., Leedham, G., 2005. Decompose algorithm for thresholding degraded historical document images. IEEE Proceedings - Vision, Image, and Signal Processing 152, 702.

[22] Drira, F., a. n.d. Towards Restoring Historic Documents Degraded Over Time. Second International Conference on Document Image Analysis for Libraries (DIAL06).

[23] Apostolos Antonacopoulos, Andy Downton, Special issue on the analysis of historical documents, International Journal on Document Analysis and Recognition 9 (April (2)) (2007) 75–77.

[24] Vijayakumar, T. (2019). Comparative study of the capsule neural network in various applications. Journal of Artificial Intelligence, 1(01), 19-27.