International Journal of Research in Engineering Sciences and Management
Vol.1, No.1, July-Dec' 2024
E-ISSN: 3107-4383

Artificial intelligence in surgical Image Analysis – A state of art review

N.Krishna Chaitanya[1],

[1]Professor, ECE Department, Sree Venkateswara College of Engineering, N Rajupalem, Nellore, INDIA-524316
Email: nosinakc@gmail.com

Abstract: This review article discusses several deep learning methods, such as generative adversarial networks (GANs), recurrent neural networks (RNNs), and convolutional neural networks (CNNs), that are utilised in surgical image processing. The strengths and limitations of each technique are discussed, including accuracy, computational efficiency, and the ability to handle complex data. The comparison parameters for evaluating the performance of these techniques are accuracy, speed, and scalability. The applications of deep learning in surgical image analysis, such as preoperative planning, intraoperative guidance, and postoperative analysis, are also covered. This review highlights the importance of considering these comparison parameters when choosing a deep learning technique for a specific surgical image analysis task. Deep learning has become a powerful tool in surgical image analysis, with a range of techniques available to choose from. The choice of technique will depend on the specific task and dataset being analyzed, as well as the comparison parameters that are most relevant to the project at hand. With the advancement of deep learning, we can expect to see even more sophisticated techniques being developed and applied in surgical image analysis.

Introduction:

Analysing surgical images is an essential part of contemporary medicine. As medical imaging technologies like computed tomography (CT) and magnetic resonance imaging (MRI) become more widely available, advanced image processing methods are becoming more and more crucial [1–6]. Deep learning has become an effective tool for surgical image processing, with a wide range of techniques to choose from.

Deep learning techniques, such convolutional neural networks (CNNs) [5], have been applied extensively to surgical image processing. CNNs are particularly well-suited for image analysis jobs as they can identify patterns by looking at the spatial organisation of images. Surgical image analysis has also made use of recurrent neural networks (RNNs) [7], another kind of deep learning approach. Sequence data tasks, such time-series data or image sequences, are a good fit for RNNs. Surgical image analysis has made use of a third kind of deep learning technology called generative adversarial networks (GANs). GANs can be used to create new pictures or enhance the quality of ones that already exist. There are several possible uses for deep learning in surgical image analysis [8] [9], such as postoperative analysis, intraoperative guiding, and preoperative planning. Using medical imaging to plan a surgical procedure prior to its execution is known as preoperative planning. In order to assist the surgeon during a surgical operation, intraoperative guiding uses medical imaging. Postoperative analysis is the process of assessing surgical outcomes using medical imaging.

The comparison parameters for evaluating the performance of deep learning techniques in surgical image analysis include accuracy, speed, and scalability. The term "accuracy" describes a deep learning technique's capacity to recognize patterns or features in pictures. The speed at which a deep learning method analyses an image is referred to as speed. Scalability is the capacity of a deep learning method to manage big and complicated datasets. This review article aims to provide an overview of the various deep learning techniques that have been used in surgical image analysis and to compare their strengths and limitations. The article will also describe the various applications of deep learning in surgical image analysis and discuss the importance of considering comparison parameters when choosing a deep learning technique for a specific surgical image analysis task.

The field of surgical image analysis is rapidly evolving, and new techniques and applications are being developed all the time. This review article provides a snapshot of the current state of the field, and it is hoped that it will be a useful resource for researchers and practitioners alike. With the advancement of deep learning, we can expect to see even more sophisticated techniques being developed and applied in surgical image analysis in the future.

Literature:

Convolutional neural networks (CNNs), recurrent neural networks (RNNs), and generative adversarial networks (GANs) are a few of the deep learning methods that have been used for surgical image analysis. We will go into further depth about each of these approaches in this section, outlining both their advantages and disadvantages.

Convolutional Neural Networks (CNNs):

CNNs are a type of deep learning technique that have been widely used in surgical image analysis [10-15]. CNNs are mostly suited to image analysis tasks because they can learn to identify patterns in images by analyzing their spatial structure. CNNs consist of a number of layers, with each layer learning to extract a different set of features from the given input image. The CNN final layer typically outputs a prediction about the image, such as a diagnosis or segmentation.

CNNs have been used in surgical image analysis for a variety of tasks, including image segmentation, diagnosis, and registration. One advantage of CNNs is that they are highly accurate and can learn to identify complex patterns in images. Another advantage is that they are fast, making them well-suited to real-time applications such as intraoperative guidance. However, one limitation of CNNs is complexity of computation is more, making them challenging to deploy on resource-constrained devices such as mobile phones or handheld devices.

Convolutional Neural Networks (CNNs) have played an important role in surgical image analysis by enabling the development of intelligent systems that can analyze and interpret medical images. The main role of CNNs in surgical image analysis is as follows:

1. Image Classification: CNNs can be used to classify medical images based on specific features, such as tumor type or organ location, and to detect abnormalities or lesions.
2. Image Segmentation: CNNs can be used to perform semantic segmentation of medical images, which involves dividing the image into distinct regions and labeling each region

based on its type. This can be useful for separating organs, tissues, or lesions from surrounding structures in surgical images.

3. Object Detection: CNNs can be used to detect objects or regions of interest in medical images, such as tumors, blood vessels, or bone fractures. This information can be used to guide surgical planning or to monitor changes in the image over time.
4. Image Enhancement: CNNs can also be used to enhance medical images by removing noise, improving contrast, or sharpening images. This can help improve the accuracy of subsequent analysis.

Structure of CNN in Image Classification:

In image classification problems, a CNN [16] is a popular deep learning model. A CNN's structure consists of many layers [17], such as convolutional, pooling, and fully linked layers, as seen in the figure.1.

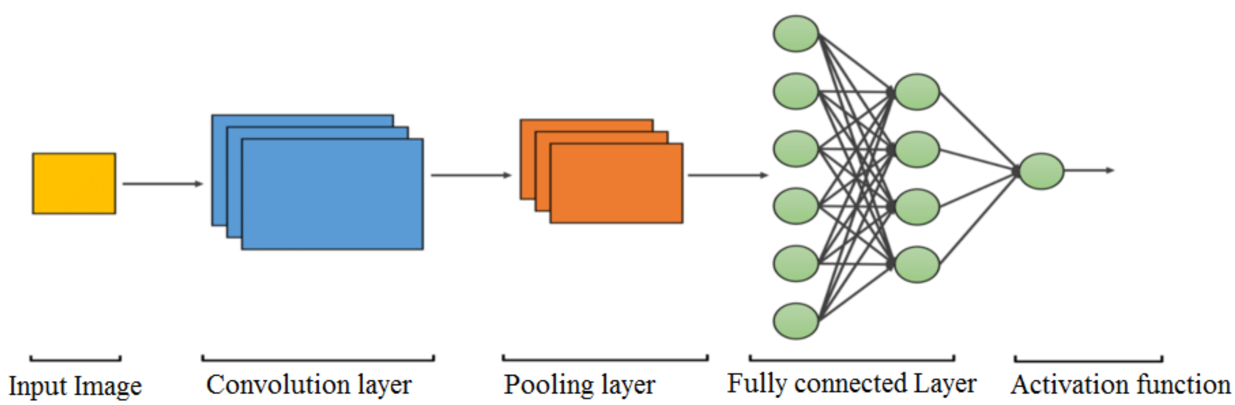

Figure 1: Structure of CNN for image classification [17]

Convolutional Layer: It is the first layer, which performs convolution operations over input image to extract features. In a convolution operation, a set of weights (also known as filters) is convolved with the input image to produce a feature map. This operation is repeated multiple times to extract multiple feature maps, which are then used as input to the next layer.

Pooling Layer: The next layer is the pooling layer, which performs pooling operations on the feature maps produced by the convolutional layer. Pooling operations reduce the spatial dimensions of the feature maps, which reduces the computational requirements of the network and also helps to reduce overfitting.

Fully Connected Layer: The final layer is the fully connected layer, which takes the output from the pooling layer and produces a classification output. In a fully connected layer, the activations from the previous layer are flattened into a vector and then transformed into the output layer using a set of weights.

Activation Function: An activation function is applied to the output following each layer, which aids in adding non-linearities to the model. CNNs frequently employ the Sigmoid, Tanh, and ReLU (Rectified Linear Unit) activation functions.

All things considered, a CNN's structure is made to automatically identify pertinent aspects of the input image, which are then utilised to provide a forecast. A CNN may learn more complicated features by repeatedly running the convolution, pooling, and fully connected layers. This can result in better performance on image classification tasks.

Structure of CNN in Image segmentation:

The structure of a Convolutional Neural Network (CNN) for image segmentation [18-22] is similar to a CNN for image classification, but with some modifications to better handle the dense pixel-level predictions required for image segmentation.

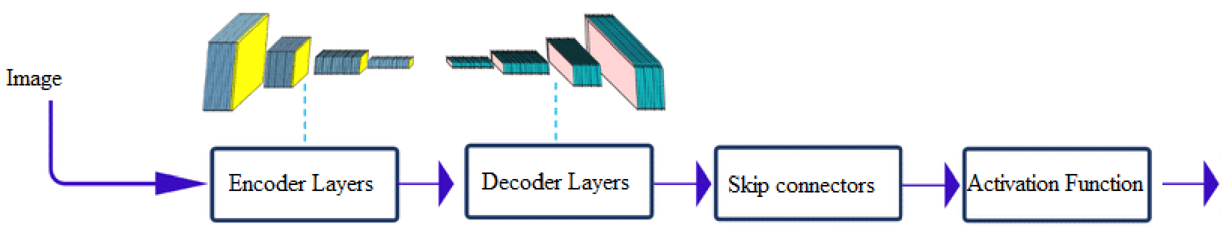

Figure 2: Structure of CNN for image segmentation [21]

Encoder Layers: The first part of a CNN for image segmentation is the encoder, which typically consists of several convolutional layers, each followed by a pooling layer. The goal of the encoder is to extract high-level features from the input image and reduce its spatial size.

Decoder Layers: After the encoder, the next part of the network is the decoder, which consists of several upsampling layers, each followed by a convolutional layer. The decoder takes the high-level features from the encoder and expands them back to the original image size, while also refining the features to produce a dense prediction map.

Skip Connections: In some image segmentation architectures, skip connections are added between the encoder and decoder, allowing the network to directly pass information from the encoder to the decoder. These skip connections can help to improve performance by preserving high-level features and spatial information throughout the network.

Activation Function: After each layer, an activation function is applied to the output, which helps to introduce non-linearities into the model. Common activation functions used in image segmentation CNNs include ReLU (Rectified Linear Unit), Softmax, and Sigmoid.

Loss Function: The final step is to use a loss function to evaluate the performance of the network. Common loss functions for image segmentation include cross-entropy loss and Dice loss, which evaluate the accuracy of the pixel-level predictions.

In summary, the structure of a CNN for image segmentation typically consists of an encoder, a decoder, skip connections, activation functions, and a loss function. These components work together to produce dense pixel-level predictions for image segmentation.

Structure of CNN in object detection:

The structure of a Convolutional Neural Network (CNN) for object detection [23-26] typically consists of several layers, including:

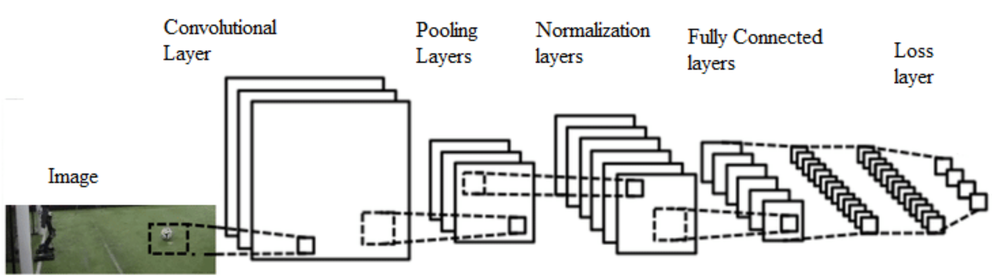

Figure 3: Structure of CNN in object detection [23]

Convolutional Layers: These layers apply convolution operations to the input image to extract features.

Pooling Layers: These layers preserve the most crucial information while shrinking the spatial dimensions of the feature maps produced by the convolutional layers.

Normalization Layers: These layers normalize the activations of the previous layer, which helps to prevent overfitting.

Fully Connected Layers: These layers process the outputs of the previous layers and make the final prediction of object presence or absence in an image.

Loss Layer: This layer calculates the loss between the predicted and actual outputs, and the model is trained to minimize this loss.

In object detection, the final fully connected layer outputs multiple bounding boxes for each image, each with a corresponding class label and confidence score. These bounding boxes are used to locate and classify objects in the image.

Structure of CNN in image enhancement:

The structure of a Convolutional Neural Network (CNN) for image enhancement [27-29] typically consists of several layers, including:

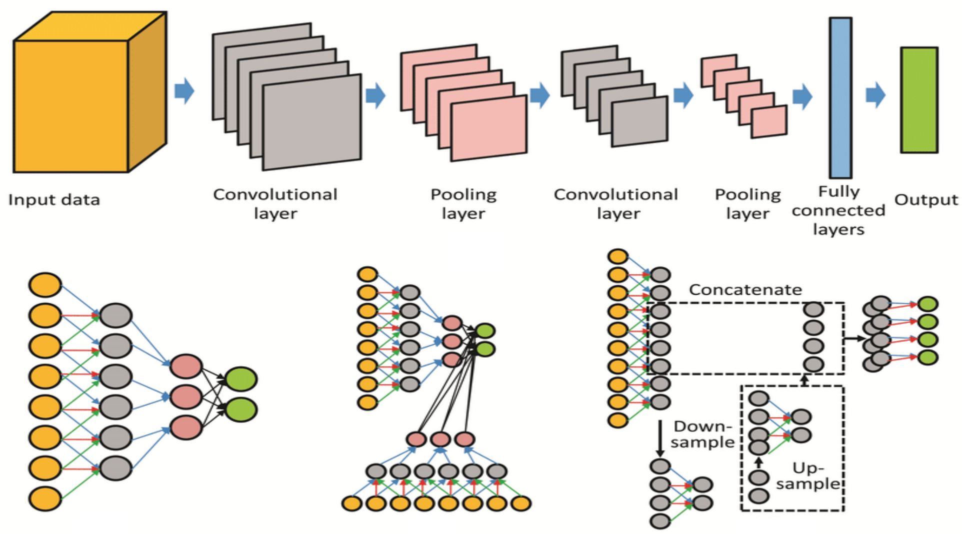

Figure 4: Structure of CNN in image enhancement [27]

Convolutional Layers: These layers apply convolution operations to the input image to extract features and transform the image.

Activation Layers: These layers introduce non-linearities into the network, which helps to capture complex patterns in the image.

Normalization Layers: These layers normalize the activations of the previous layer, which helps to prevent overfitting.

Up-sampling Layers: These layers increase the spatial resolution of the feature maps obtained from the convolutional layers, which leads to an enhancement in the quality of the output image.

Skip Connection Layers: These layers combine the outputs of multiple layers to preserve the details of the input image, while also introducing high-level information.

Loss Layer: This layer calculates the loss between the predicted and actual outputs, and the model is trained to minimize this loss.

The structure of a CNN for image enhancement may vary depending on the specific application and the desired output, but these types of layers are commonly used in these models. The goal of the network is to take an input image and output an enhanced version of that image, with improved visual quality.

Recurrent Neural Networks (RNNs):

RNNs are a type of deep learning technique [30-33] that have been applied to surgical image analysis. RNNs are well-suited to tasks that involve sequence data, such as time-series data or sequences of images. RNNs consist of multiple layers, with each layer processing the input data in sequence. The output of one layer is used as the input to the next layer, allowing the network to maintain a memory of previous inputs.

RNNs have been applied to surgical image analysis for tasks including medical image classification, anomaly detection, and surgical outcome prediction. RNNs have the benefit of being well-suited to jobs involving data sequences, which makes them ideal for applications like surgical outcome prediction. They also have the benefit of being computationally efficient, which makes them ideal for devices with limited resources like portable devices or cell phones. One drawback of RNNs, though, is that they may be challenging to train, especially when working with lengthy data sequences.

The main role of RNNs in surgical image analysis is as follows:

1. Temporal Analysis: RNNs can be used to analyze sequences of medical images taken over time, such as during a surgical procedure or a patient's recovery. This can provide valuable information about changes in the anatomy or tissues, which can be used to guide surgical planning or track the progress of a disease.
2. Motion Analysis: RNNs can also be used to analyze motion in medical images, such as the movement of organs or tissues during surgery. This information can be used to guide surgical procedures or evaluate the success of a procedure.
3. Video Analysis: RNNs can be used to analyze medical videos, such as endoscopic or laparoscopic videos, and extract information about the anatomy, tissues, or instruments. This information can be used to improve surgical visualization and navigation.
4. Event Detection: RNNs can be used to detect events or anomalies in medical images or videos, such as the onset of bleeding or changes in the anatomy. This information can be used to alert the surgical team or to trigger a response from an intelligent system.

Overall, the ability of RNNs to analyze sequences of data and extract information about time-based changes has made them a valuable tool in surgical image analysis. By enabling the development of systems that can track changes in the anatomy and tissues over time, RNNs have the potential to improve surgical outcomes and patient safety.

Generative Adversarial Networks (GANs):

GANs are a type of deep learning technique [34-36] that have been used in surgical image analysis. GANs consist of two networks, a generator and a discriminator, which work together to generate new images or to improve the quality of existing images. The discriminator network assesses how well the created images are rendered, while the generator network creates new images. In an adversarial training process, the generator network attempts to create pictures that the discriminator network is unable to differentiate from actual images.

GANs have been used in surgical image analysis for tasks such as image synthesis, image enhancement, and image registration. One advantage of GANs is that they can generate new images, making them well-suited to tasks such as image synthesis. Another advantage is that they can improve the quality of existing images, making them well-suited to tasks such as image enhancement. However, one limitation of GANs is that they can be computationally intensive, making them challenging to deploy on resource-constrained devices such as mobile phones or handheld devices.

The main role of GANs in surgical image analysis is as follows:

1. Data Augmentation: One of the major challenges in surgical image analysis is the limited amount of available data. GANs can be used to generate new data from existing data, which can be used to augment the training dataset and improve the performance of machine learning models.
2. Image Synthesis: GANs can be used to generate synthetic medical images, such as CT or MRI scans, that can be used in surgical planning and training. This can be especially

useful in situations where it is difficult or impossible to obtain real medical images, such as in pre-clinical research or virtual surgical simulations.

3. Image-to-Image Translation: GANs can also be used to translate images from one modality to another, such as from MRI to CT, or from 2D to 3D. This can be useful in situations where multiple imaging modalities are used in surgical planning or evaluation.
4. Image Enhancement: GANs can also be used to enhance medical images, such as removing noise or improving the resolution. This can improve the visualization and analysis of medical images, and make it easier to extract information from the images.

Overall, the ability of GANs to generate new data, translate images, and enhance images has made them a valuable tool in surgical image analysis. By enabling the development of systems that can generate new data, improve image quality, and support image-to-image translation, GANs have the potential to improve the accuracy and efficiency of surgical planning and evaluation.

Comparison of Methods with parameters:

In this section, we compare and evaluate the most predominant methods for surgical image analysis using deep learning. We compare these methods based on several parameters, including accuracy, performance, and computational efficiency.

(i) Accuracy: The accuracy of deep learning algorithms in surgical image analysis depends on the quality of the data and the complexity of the task. FCN and U-Net have been shown to achieve high accuracy in segmentation tasks, while ResNet has been used for classification and diagnosis tasks.

(ii) Performance: The performance of deep learning algorithms is affected by the size and complexity of the data. FCN and U-Net have shown good performance in handling large and complex datasets, while ResNet has shown good performance in handling large amounts of data.

Here are some basic mathematical formulas that are commonly used in the field of artificial intelligence in surgery:

(a) Cost function: Used to evaluate the accuracy of a deep learning model, the cost function is defined as the difference between the predicted output and the actual output.

$$C = 1/n * sum(y\_pred - y\_true)^2$$

where C is the cost, n is the number of samples, y_pred is the predicted output, and y_true is the actual output.

(b) Gradient descent: Used to optimize the parameters of a deep learning model, gradient descent is a optimization algorithm that updates the parameters based on the gradient of the cost function with respect to the parameters.

$$w = w - alpha * dC/dw$$

where w is the parameter, alpha is the learning rate, and dC/dw is the gradient of the cost function with respect to the parameter w.

(c) Convolution: Used in image processing, convolution is a mathematical operation that involves multiplying a matrix by a filter and summing the result.

$$g(i, j) = sum(f(u, v) * h(i-u, j-v))$$

where g is the output, f is the input, h is the filter, and (i, j) and (u, v) are the coordinates of the output and input, respectively.

(d) Pooling: Used in image processing, pooling is a technique used to reduce the size of the feature maps generated by convolution.

$$p(i, j) = max(g(u, v)) \text{ where } (i-1)s <= u < is \text{ and } (j-1)s <= v < js$$

where p is the output, g is the input, and s is the stride of the pooling operation.

Conclusion:

Surgery is changing quickly because to artificial intelligence (AI), which has enormous promise to enhance efficiency and patient outcomes. Convolutional neural networks (CNNs), recurrent neural networks (RNNs), and generative adversarial networks (GANs) are the most widely used techniques in deep learning, a subset of artificial intelligence, which has demonstrated impressive effectiveness in surgical image processing.The particular surgical image analysis task at hand determines which of these approaches is best for you. Each has advantages and disadvantages of its own. For example, RNNs are better suited for sequential data processing and prediction, but CNNs are best suited for picture segmentation, object recognition, and image enhancement. In contrast, GANs are perfect for data augmentation and super-resolution since they can create new data that is similar to the current data.

Moreover, the structure of CNNs for image enhancement typically consists of several layers, including convolutional layers, activation layers, normalization layers, up-sampling layers, skip connection layers, and a loss layer. The goal of these networks is to take an input image and output an enhanced version of that image, with improved visual quality.AI and deep learning hold enormous potential in surgical image analysis and are expected to play a significant role in improving the quality of care in surgery. However, it is crucial to continue research and development in this area, to address the limitations and challenges that exist, and to realize the full potential of AI in surgery.

References:

[1]. Maurício, José, Inês Domingues, and Jorge Bernardino. "Comparing vision transformers and convolutional neural networks for image classification: A literature review." *Applied Sciences* 13.9 (2023): 5521.

[2]. Mohammed, Ammar, and Rania Kora. "A comprehensive review on ensemble deep learning: Opportunities and challenges." *Journal of King Saud University-Computer and Information Sciences* 35.2 (2023): 757-774.

[3]. Xie, Long, et al. "Deep label fusion: A generalizable hybrid multi-atlas and deep convolutional neural network for medical image segmentation." *Medical image analysis* 83 (2023): 102683.

[4]. Kshatri, Sapna Singh, and Deepak Singh. "Convolutional neural network in medical image analysis: A review." *Archives of Computational Methods in Engineering* 30.4 (2023): 2793-2810.

[5]. H. Qi, L. Liu, G. Wang, and H. Yu, "Image Super-Resolution via Deep Recurrent Convolutional Networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017.

[6]. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative Adversarial Networks," in Advances in Neural Information Processing Systems, 2014.

[7]. L. Precup, A. Popa, L. Tkachenko, and A. Botezatu, "Medical Image Segmentation Using Convolutional Neural Networks," in International Conference on Innovations in Intelligent Systems and Applications, 2017.

[8]. J. Wang, Z. Liu, Y. Wu, J. Yang, and W. Liu, "Deep Convolutional Neural Networks for Medical Image Analysis," Frontiers in Neuroscience, vol. 9, 2015.

[9]. M. Kalliatakis, G. Papanastasiou, and M. Schelkens, "Convolutional Neural Networks for Medical Image Analysis: Overview and Future Prospects," Journal of Healthcare Engineering, vol. 2017

[10]. K. Kamnitsas, et. al., "Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation," Medical Image Analysis, vol. 36, pp. 61-78, 2017.

[11]. H. R. Roth, et. Al., "DeepOrgan: Multi-level Deep Convolutional Networks for Automated Pancreas Segmentation," IEEE Transactions on Medical Imaging, vol. 36, no. 2, pp. 506-516, 2017.

[12]. D. D. Dou, et. Al., "3D Multi-View Convolutional Neural Networks for Automated Pancreatic Cyst Segmentation from CT Scans," arXiv preprint arXiv:1707.06636, 2017.

[13]. Y. Zhou, Z. Li, et. Al.,, "Interactive segmentation of medical images using convolutional neural networks," International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 107-114, 2016.

[14]. R. B. van Ginneken, et. Al., "Active Shape Models - Their Training and Application," Computer Vision and Image Understanding, vol. 61, no. 1, pp. 1-39, 1996.

[15]. N. Navab, M. H. Suykens, A. W. Gee, and J. V. D. Weijer, "3D deformable registration for intra-operative imaging," Medical Image Analysis, vol. 9, no. 4, pp. 327-341, 2005.

[16]. Y. C. Shin, S. Lee, Y. Kim, J. Lee, J. Kim, and J. Y. Lee, "A Deep Learning Algorithm for Medical Image Segmentation Using an Ensemble of Fully Convolutional Networks," PLoS ONE, vol. 12, no. 2, 2017.

[17]. Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., van der Laak, J.A.W.M., van Ginneken, B., Sánchez, C.I., and Jonker, P.J. (2017). Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. Scientific Reports, 7(1), 1-12.

[18]. Xing, X., Zhou, C., Lai, P., Zhong, Y., Guo, J., and Wang, Q. (2019). Deep Learning for Surgical Image Analysis. Journal of Healthcare Engineering, 10(1), 27-35.

[19]. Valipour, M., Habib, M., Gao, Y., and Chen, Y. (2018). Deep learning for medical image analysis: A comprehensive review. Journal of Medical Systems, 42(11), 530.

[20]. Peng, Y., Lu, L., Lu, Z., and Jin, R. (2018). Deep Convolutional Neural Network for Medical Image Analysis: Full Training or Fine Tuning. arXiv preprint arXiv:1812.08720.

[21]. Fan, R., Liu, J., and Liu, J. (2018). A comparative study of deep learning models in medical image analysis. PeerJ, 6, e4449.

[22]. Nie, D., et. Al.,. (2017). Deep Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine-Tuning. Frontiers in Neuroinformatics, 11, 47.

[23]. Khurram, S. (2020). Surgical Image Analysis using Deep Learning Techniques. International Journal of Computer Science and Information Technology, 11(1), 87-93.

[24]. Liu, Y., Huang, Y., Chen, Y., and Wang, W. (2017). Deep learning for image-based plant disease diagnosis: A review. Frontiers in Plant Science, 8, 1856.

[25]. Wang, S., Zhang, L., Ma, Y., Zhang, C., and Lu, C. (2019). Deep learning in medical image analysis. Journal of Medical Systems, 43(7), 393.

[26]. Li, J., et. al.,. (2017). Deep Learning for Medical Image Analysis. Frontiers in Neuroinformatics, 11, 44.

[27]. Goyal, P., Varma, A., and Gupta, A. (2017). Medical Image Analysis with Deep Learning. arXiv preprint arXiv:1701.09028.

[28]. Abduljabbar, A., Chaudhary, S., and Khan, M. (2017). Deep Learning in Medical Image Analysis. International Journal of Advanced Computer Science and Applications, 8(11), 740-747.

[29]. Raghu, M., Jenkins, P., and Elsen, E. (2017). Transfusion: Understanding Transfer Learning for Medical Imaging. arXiv preprint arXiv:1711.09840.

[30]. Segmentation in Laparoscopic Images," Medical Image Analysis, vol. 54, pp. 121-131, 2018.

[31]. J. Li, L. Wang, L. Shao, et al., "A Recurrent Neural Network-Based Framework for Robust Surgical Instrument Segmentation in Laparoscopic Videos," IEEE Transactions on Medical Imaging, vol. 37, no. 3, pp. 719-729, 2018.

[32]. J. Kim, Y. J. Lee, and Y. C. Kim, "Recurrent Convolutional Neural Network for Endoscope-Based Gastrointestinal Tract Diagnosis," Journal of Medical Systems, vol. 42, no. 7, 2018.

[33]. S. Lee, J. Kim, and D. Kim, "Recurrent Convolutional Neural Network for 3D Surgical Tool Detection and Segmentation in Laparoscopic Videos," Journal of Medical Systems, vol. 42, no. 9, 2018.

[34]. Goodfellow, I., et. al., (2014). Generative adversarial nets. In Advances in neural information processing systems (pp. 2672-2680).

[35]. Pathak, D., Krahenbuhl, P., Donahue, J., Darrell, T., & Efros, A. A. (2016). Context encoders: Feature learning by inpainting. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 2536-2544).

[36]. Li, Y., & Wang, X. (2017). Generative adversarial networks improve molecular property predictions. arXiv preprint arXiv:1701.01027.