Enhanced Solar Potential Analysis: Separating Terraced House Rooftops Using Convolutional Neural Networks

Kai Zhang*
Department of Computing
Goldsmiths, University of London
London, United Kingdom
kzhan002@gold.ac.uk

Dr. Lahcen Ouarbya
Department of Computing
Goldsmiths, University of London
London, United Kingdom
l.ouarbya@gold.ac.uk

***Abstract*—Solar power, a clean and renewable energy source, plays a pivotal role in achieving sustainable development goals by offering affordable, reliable, modern energy solutions and mitigating energy-related emissions and pollutants. Current studies predominantly focus on solar potential analysis derived from machine learning-based rooftop area segmentation. However, these studies reveal an overestimation of usable area for solar output calculations in terraced houses, due to failing to distinguish individual households within terraced structures. This research delineates state-of-the-art Machine Learning and computer vision techniques applied on remote-sensing images obtained via the Google API [1]. The dataset, manually annotated and augmented to include 5000 training images and 1000 validation images, is focused on the UK, particularly terraced house areas. The standalone Convolutional Neural Network used to segment terraced-structure rooftop areas reaches an intersection over union of 69.11%. The model uniquely addresses the segmentation of contiguous terraced houses in the UK, which is pivotal for the solar installation assessments in the UK's residential landscape.**

Keywords: Solar Potential, Renewable Energy, Roof Segmentation, Terraced Houses, Convolutional Neural Network, Remote Sensing

I. Introduction

Solar energy is one of the most abundant and clean alternatives to conventional hydrocarbon fuels, and one of the most promising for facilitating global access to economic, reliable, and secure energy.

In 2015, the United Nations (UN) launched an agenda encompassing 17 Sustainable Development Goals (SDGs) aimed at fostering global sustainable development and prosperity [2]. Solar energy, as a renewable source with reduced carbon emissions compared to fossil fuels, holds significant promise in meeting the escalating energy demands, thanks to advancements in technology and cost reductions.

While numerous studies and tools have been developed to assess potential solar output on specific rooftops, a significant lack arises in countries like the UK, where terraced houses constitute a major portion of housing. 26.3% [3]. A terrace, terraced house (UK), or townhouse (US) is a kind of medium-density housing built as part of a continuous row in a uniform style.

Existing machine learning models missed out this structural fact leading to a lack of segment individual rooftops of terraced houses. This oversight potentially leads to underutilization of solar panels due to decreased household interest as the inaccessibility of solar assessment of their own rooftops. By resolving this issue, our approach not only provides solar potential estimates for a broader range of housing types but also supports the UK's ambitious target of achieving 40 gigawatts of solar generation capacity by 2030[4]. In the next section a detailed discussion of existing studies in this area is given, what are their strengths, as well as why segmenting terraced houses to individual level is important for current machine learning driven tool to assess solar potential of rooftops. Figure 1 below is a typical type of terraced houses in the UK. An example satellite image is given in Figure 6.

Fig. 1: UK Terraced Houses

A. Existing Studies

Numerous studies have been undertaken to analyze the potential of rooftop solar installations. Izquierdo et al. (2008)[5] initiated this research trajectory by assessing available rooftop areas, a critical preliminary step in understanding the spatial capacity for solar installations. Subsequent research expanded on this by incorporating detailed building characteristics to estimate solar potential from a bottom-up perspective, as demonstrated by the works of Korfiati et al. (2016)[6].

A notable advancement in this domain is the introduction of the SolarNet framework by Qingyu Li (2023)[7], which utilized CNNs for comprehensive rooftop segmentation and solar potential analysis from aerial imagery. This framework proposes a simplified representation of roof geometries, reducing the number of classes to mitigate class imbalance and ease annotation efforts, thus streamlining the solar potential analysis process.

Recent advancements in the field of rooftop solar potential analysis have leveraged high-resolution imagery and sophisticated machine learning algorithms to enhance accuracy and efficiency. A notable study utilized object-based image classification combined with the normalized difference vegetation index (NDVI) and digital surface models (DSMs) to estimate rooftop solar photovoltaic potential accurately. This method was validated using high-resolution imagery from Beijing, demonstrating its effectiveness in identifying suitable rooftops for solar panel installation and estimating their photovoltaic potential with considerable precision[8].

Another innovative approach employed a two-stage process for estimating solar photovoltaic potential on building rooftops. Initially, rooftops were detected using satellite images processed through a series of image pre-processing algorithms alongside machine learning techniques, specifically Support Vector Machine (SVM) and Naïve Bayes (NB). The subsequent stage of solar PV potential estimation utilized tools like the PVWatts calculator, PVGIS, and ArcGIS. Applied to the B6 division of Madinaty City in Egypt, this method achieved high metrics in precision, recall, and F1-score for rooftop detection and provided robust estimates of the annual photovoltaic potential, underscoring the method's utility in supporting sustainable energy initiatives[9].

In the specific realm of rooftop area estimation, methodologies utilizing aerial imagery have emerged as a prominent approach due to their ability to provide current and high-resolution data. Among these, Lee et al. (2019)[10], Huang et al. (2019)[11] has made significant strides in roof segmentation using aerial images. However, while effective in delineating roof boundaries, these methods often neglect the intricate details like roof tilts and orientations, which are pivotal for precise solar potential calculations.

Lee et al.[10] partially addressed this challenge by employing convolutional neural networks (CNNs) to classify roof segments according to their orientations, a method that, while innovative, faced difficulties with complex annotation requirements and class imbalance issues. These challenges were further tackled by Krapf et al.[12] by proving a roof information dataset which can be leveraged as the dataset for CNNs for roof segmentation. However, Krapf also pointed out the risks of biased network learning due to imbalanced data. Rodrigo et al.[13] has taken the advantage of the dataset provided by Krapf et al.[12], which was collected in Wartenberg, Germany, to train a roof segmentation CNNS model with an intersection over union (IoU) 0.49.

B. The problem remains unsolved

Rooftop solar installations represent a significant frontier in the quest for sustainable energy solutions. The accuracy of rooftop detection for solar panel placement is crucial for maximizing energy efficiency and viability. However, the challenge of accurately assessing all various rooftop types with varied rooftop configurations prevail.

As indicated in Fig.3, current methodologies predominantly approach rows of terraced houses as singular entities, thereby neglecting the necessity of segmenting individual households. This oversight is particularly significant in regions such as the United Kingdom, where terraced housing constitutes a major portion of residential structures [3]. Such a generalized approach can lead to substantial inaccuracies in estimating the solar output potential for a vast number of residences.

The literature reveals a gap in the segmentation accuracy of these models, particularly regarding their application in the nuanced context of terraced housing prevalent in regions like the UK. For example, studies have often treated contiguous rooftop arrays of terraced houses as single units, thereby skewing potential solar capacity estimates [7].

Fig. 2: Scattered houses solar potential(Yellow bounding box) and terraced houses overestimated solar potential(Red bounding circle) taken from [7]

This mistreatment of terraced houses rooftops potentially compromises the motivation of households of terraced houses from converting to solar energy due to the lack of accessibility to their individaulised rooftop solar potential estimate drivied from exising solar potential estimating tools such as Project Sunroof from Google[14]. As reveals in figure 3, any one of the three households(bounded by green boxes) in this particular terraced house was not able to get an estimate of how much solar energy its own rooftop can produce(bounded by cyan boxes) if they decide to install solar panels from existing tools. This is because the exiting studies and tools treat the three connected terraced houses as one(bounded by red box)which could potenrially leads to a compromised motivation to install solar panels on their rooftops.

Fig. 3: A example of terraced houses seperation problem

Rodrigo et al.[13]has conducted a study dedicated in 2023 to resolved this issue discussed above, however, due to the imbalanced distribution of the dataset used, a relatively low IoU of 0.49 was achieved.

The study proposed in this paper seeks to address the gap discussed in Section I-B and provides an enhanced assessments of solar energy potential for individual terraced households by leveraging a CNN machine learning model trained on a more enhanced dataset. Next session discusses the data collection and data processing of the data using in this study.

II. Data Collection and Pre-processing

In the realm of semantic segmentation of roof segments, the literature primarily cites two benchmark datasets: the DeepRoof dataset [10] and the RID dataset [12]. DeepRoof encompasses annotations for 2,274 buildings, while RID features 1,880 annotated buildings. Notably, both datasets predominantly represent data outside the UK and incorporate a diverse array of universal roof types.

To address these limitations and focus specifically on UK terraced roofs, we have developed a tailored validation dataset. This dataset was meticulously constructed through manual annotation of satellite images sourced from Google Maps Platform API[1], specifically images showcasing terraced rooftops with discernible "gaps" between individual households. The annotation process was conducted using the Computer Vision Annotation Tool (CVAT) [15] .

To enhance the dataset's utility and representativeness, we employed image augmentation techniques[16], effectively scaling the dataset to approximately 2,000 images. First, we randomly flip it horizontally, vertically or rotate it by 90 degrees: this increased by eight times the dimension of our dataset. Second, we randomly crop each image to a 248x248 size, which helps in detecting rooftops lying at the edges of the image. Third, we randomly add uniform and Gaussian noise to avoid model overfitting[17]. This augmentation ensures a more robust and varied dataset, facilitating more comprehensive model training and validation.

It is crucial to note that our study is distinctly focused on the segmentation of connected rooftop structures, such as those found in terraced houses. Consequently, our validation set is deliberately designed to exclude other aspects of semantic roof segmentation, such as the geometric configuration of rooftops and the orientation of individual roof segments. This focused approach is in line with our specific research objectives and helps in refining the accuracy of our segmentation model for the targeted application of terraced house roof separation. This proposed methodology is discussed in the next section.

III. Methodology

Convolutional Neural Network : We adapt the U-Net architecture[18], a CNN developed for biomedical images segmentation, to our task and inputs. The network originally comes with two parts. A contracting path (encoder) extracts features at different levels through a sequence of convolutions, filters, activation functions and pooling layers, allowing to capture the context of each pixel. Then a symmetric expanding path (decoder) upsamples the results, increasing the resolution of the detected features. The output of each stage of the down-sampling phase is fed directly to the corresponding up-sampling phase to avoid separate training of encoder and decoder. U-Net has shown great performances with only a limited number of training images. In order to detect the available rooftop area to install RPV modules, we slightly modify the original architecture. We restrict the output of the model to a pixel-wise binary classification (suitable rooftop area or not suitable) and we adapt the number of filters and the input size in order to match the 512x512 pixels. Figure 1a shows the modified architecture, which has in total 14'788'929 parameters. For each node we choose the Rectified Linear Unit activation function followed by batch normalisation. Since we are interested in the segmented area of an image, the output vector has a sigmoid activation function, which sets for each pixel the probability that it belongs to an area in the image suitable for RPV installation (positive class). Each pixel is then attributed to the positive class when its value exceeds a probability threshold, which is set at 50%.

Model training :The whole dataset is split into three sets: 80% of images for the training, 10% for validation, and 10% for testing. We train the U-Net architecture starting from a random set of weights. The model is trained on small batches of two images at a time. Since the goal is to classify each pixel either as available or not for solar RPV installation, the binary cross entropy loss is the natural choice. The proportion of pixels that are labelled as available is low compared to non-available ones, creating an unequal distribution of the two classes. We overcome this potential bias by applying weighted binary cross entropy loss, setting the weight for the less frequent class to 4. Accuracy is the standard metric for classification tasks. Given the uneven class frequency in images, we apply Intersection Over Union (IoU or Jaccard Index) discussed in Section IV, which is a more suited metric for unbalanced datasets. We rely on IoU in order to evaluate the training performance of our model. In the literature an IoU larger than 0.5 is considered a good prediction. As gradient-based optimization algorithm, we use Adam[19], commonly known in computer vision tasks to speed up the convergence, setting the default first and second moment estimates to 0.9 and 0.999. We start from a learning rate equal to 0.0008 and we further reduce it by a factor 0.8 every 10 epochs. An example mask produced by the model trained on learning rate of 0.0008 is shown in Figure 4As observed in Figure 5, 10 epochs are sufficient to allow the IoU on the validation set to converge. The final metrics obtained on training, validation and test set are reported in Table I in in Section V.

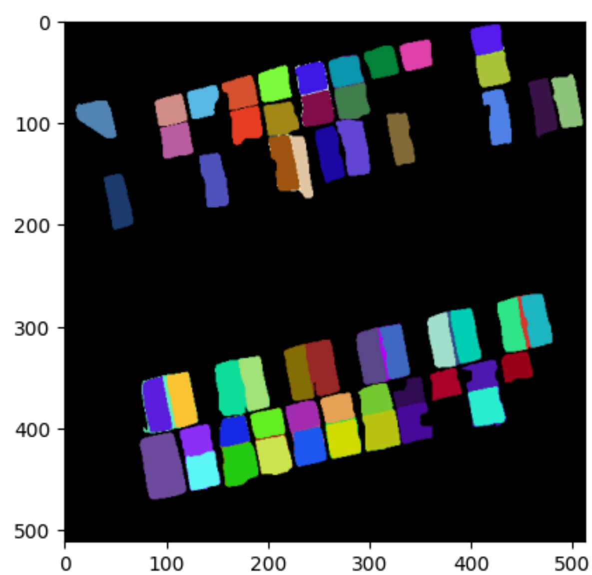

Fig. 4: Example mask with 0.0008 learning rate after color processing

A. Technical Details

The **U-Net** model, originally conceived for biomedical image segmentation[18], is adept at capturing multi-scale features crucial for delineating complex rooftop geometries. This paper extends this model within the proposed framework, leveraging its encoder-decoder architecture for robust feature

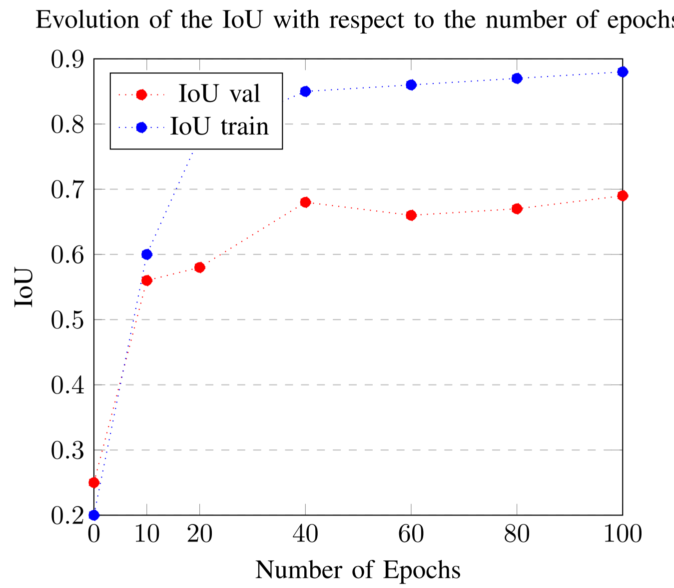

Fig. 5: Evolution of the IoU with respect to the number of epochs

extraction and contextual interpretation. Specifically, the U-Net's encoder component, derived from a VGG-16 backbone, is employed for hierarchical feature extraction, characterized by the encoding sequence given in Equation 1.

$$E(x) = \{e_1(x), e_2(x), \cdots, e_(x)\} \tag{1}$$

where $E(x)$ represents the encoded feature set for an input image x, and ei(x) denotes the feature maps at encoding layer i. These feature maps are subsequently utilized in the decoder phase to reconstruct the segmented rooftop profile with enhanced spatial resolution given in Equation 2 below:

$$D(E(x)) = d_n(e_n(\cdots(d_2(e_2(d_1))\cdots)) \tag{2}$$

where $D(E)$ signifies the output from the decoder, leveraging the contextual and spatial information encapsulated within $E(x)$.

The integration with UNetPlusPlus, a sophisticated iteration of the U-Net model[20], introduces an augmented set of skip pathways, enhancing feature propagation and reducing semantic gaps across the encoder-decoder interface. This modification is paramount in refining the model's sensitivity to rooftop edges and textures, facilitating a more granular segmentation output.

IV. Validation

The validation process is a critical component in assessing the performance of our terraced house roof segmentation framework. We employ a combination of established and innovative validation methodologies to ensure the robustness and accuracy of our model.

- **Metric - Mean Intersection Over Union (mIOU):** mIOU is a pivotal metric in our validation process. It quantifies the overlap between our model's predictions and the ground truth annotations, providing a reliable measure of segmentation accuracy. The use of mIOU is crucial for objectively evaluating the model's performance across various scenarios.
- **Annotation and Dataset Creation:** The creation of a high-quality validation dataset is foundational to our validation strategy. Utilizing the Computer Vision Annotation Tool (CVAT)[15] ensures precise and consistent annotations. This meticulous annotation process lays the groundwork for a valid comparison between the model's predictions and the ground truth.
- **Data Source and Diversity:** Our validation dataset is sourced from a wide range of satellite images across the UK, encompassing diverse terraced house roof types and urban landscapes. This diversity is key to evaluating the model's generalizability and effectiveness in different real-world scenarios.
- **Validation Procedure:** The validation involves a thorough comparison of the segmented outputs against the manually annotated ground truth. The computation of mIOU scores for each image in the dataset offers a detailed assessment of segmentation precision.

V. Results and Discussion

The CNN-based model presented in this study is validated as discssed in the last section. The proposed model has achieved an Intersection over Union (IoU) 69.11%, as shown in Table I. This enhancement is particularly significant in the context of solar potential analysis, where precision in identifying and segmenting rooftops, especially in complex urban landscapes like terraced housing, is crucial. The model effectively addresses the challenge of potential compromised motivation of converting to solar energy due to mistreatment of terraced houses in the existing studies by providing a separated rooftop detection for terraced houses from remote imagery, Figure 9 is an example mask of successful case of our model separating terraced house rooftops. This mask was processed by our model on Figure 6, the colored area in the Figure 7 is the targeted rooftops waiting to be separated. Figure 8 is model mask overlaid on Figure 6.

In short, The highlighted areas in Figure 7 are the wanted separating area. After being processed by our model on Figure 6, Figure 9 is produced, each colored segment represents an individual household in the targeted separating areas.

TABLE I: Performance of the CNN evaluated on the training, validation and test sets.

	IoU	Accuracy	Recall	Precision
Training	0.8823	0.9794	0.9299	0.9437
Validation	0.6911	0.9264	0.8360	0.8508
Test	0.6420	0.9307	0.7522	0.7874

Fig. 6: An example terraced houses satellite image in the UK

Fig. 7: Targeted houses to separate

The Table II below illustrates the IoU performance comparison of existing studies in rooftop segmentation in segments and our model:

Study	IoU Performance (%)	Dataset Region
Krapf et al.[12]	49	Wartenberg, Germany
Our Model	69	UK

TABLE II: IoU performance comparison across various studies

The main contribution of this paper compare to other existing studies is to introduce a model to address the issue

Fig. 8: Separated individual households

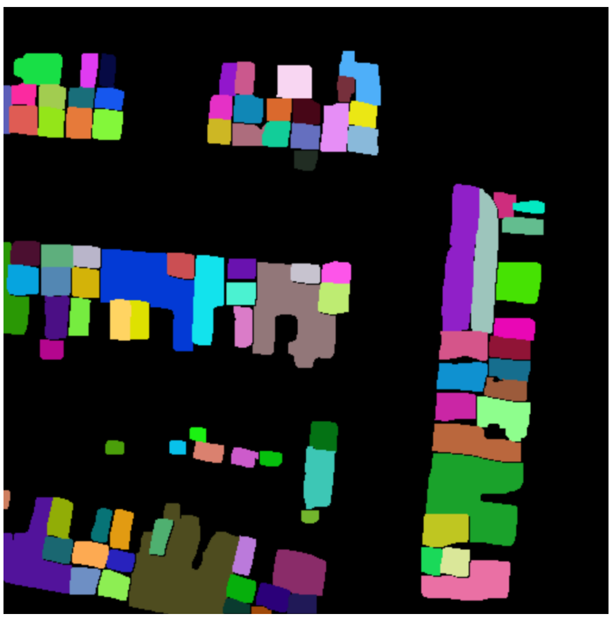

Fig. 9: The mask of separated individual households

of individual roof segmentation of terraced houses, which increases the accessibility to the household of terraced houses to an individual level, while maintain the high accuracy of detection which has been achieved by other studies already. By providing this enhanced individual level segmentation of terraced houses, a larger portion of terraced house household pivoting to solar energy is expected.

VI. Conclusion

In this work we present a novel method to estimate the available rooftop surfaces for terraced houses by using state-of-the-art ML and computer vision techniques based on aerial images. The model is built on a dataset that distinctly focused on the segmentation of connected rooftop structures, such as those found in terraced houses. The model is validated on a dataset designed to exclude other aspects of semantic roof segmentation, such as the geometric configuration of rooftops and the orientation of individual roof segments. Based on this, the available rooftop areas of terraced houses are detected with an IoU of 69% and an accuracy of 92%. Assuming this CNN model based on aerial images yields realistic results, we compare it with an existing large-scale estimate of available roof area, our model is able to separate rooftops of row-structure houses such as terraced houses while maintain a same level of accuracy. Such an estimate at the resolution of individual roofs may provide valuable insights for the real world application to provide an individual-level solar estimate to a large portal houses in the counties such as UK, potentially contributes to the UK government to achieve its goal to reach to a solar generation capacity by 2030.

References

[1] *Google static map*, https://developers.google.com/maps/documentation/maps-static.

[2] *Google project sunroof*, https://www.europarl.europa.eu/RegData/etudes/IDAN/2022/702576/EXPO_IDA(2022)702576_EN.pdf.

[3] *Official statistics: Council tax: Stock of properties statistical summary perties by property type*, https://www.gov.uk/government/statistics/council - tax - stock - of - properties - 2021 / council - tax - stock - of - properties-statistical-summary.

[4] *Written evidence submitted by solar energy uk*, https://committees.parliament.uk/writtenevidence.

[5] M. R. Salvador Izquierdo and N. Fueyo, "A method for estimating the geographical distribution of the available roof surface area for large-scale photovoltaic energy-potential evaluations," *Journal of Solar Energy*, vol. 82, 2098.

[6] F. V. e. a. Athina Korfiati Charalos Gkonos, "Estimation of the global solar energy potential and photovoltaic cost with the use of open data," *International Journal of Sustainable Energy Planning and Management*, 2016.

[7] Q. Li, S. Krapf, Y. Shi, and X. X. Zhu, "Solarnet: A convolutional neural network-based framework for rooftop solar potential estimation from aerial imagery," *International Journal of Applied Earth Observation and Geoinformation*, vol. 116, p. 103 098, 2023, ISSN: 1569-8432. DOI: https://doi.org/10.1016/j.jag.2022.103098.

[8] B. HuMengjin ZhaoLiu, "Estimation of rooftop solar photovoltaic potential based on high-resolution images and digital surface models," *Buildings*, 2023.

[9] S. M. Eslam Muhammed Adel El-Shazly, "Building rooftop extraction using machine learning algorithms for solar photovoltaic potential estimation," *Sustainability*, 2023.

[10] S. Lee, S. Iyengar, M. Feng, P. Shenoy, and S. Maji, "Deeproof: A data-driven approach for solar potential estimation using rooftop imagery," *25th ACM SIGKDD International Conference on Knowledge Discovery Data Mining*, DOI: 10.1145/3292500.3330741.

[11] X. S. Huang Z. Mendis T., "Urban solar utilization potential mapping via deep learning technology: A case study of wuhan, china," *Applied Energy*, 2019.

[12] S. Krapf, L. Bogenrieder, F. Netzler, G. Balke, and M. Lienkamp, "Ridmdash;roof information dataset for computer vision-based photovoltaic potential assessment," *Remote Sensing*, vol. 14, no. 10, 2022, ISSN: 2072-4292. DOI: 10.3390/rs14102299.

[13] R. Pueblas, P. Kuckertz, J. M. Weinand, L. Kotzur, and D. Stolten, "Ethos.passion: An open-source workflow for rooftop photovoltaic potential assessments from satellite imagery," *Solar Energy*, vol. 265, p. 112 094, 2023, ISSN: 0038-092X. DOI: 10.1016/j.solener.2023.112094.

[14] *Google project sunroof*, https://sunroof.withgoogle.com/.

[15] CVAT.ai Corporation, *Computer Vision Annotation Tool (CVAT)*, version 2.8.2, Nov. 2023.

[16] S. S. T. M. Khoshgoftaar, *A survey on image data augmentation for deep learning*, https://doi.org/10.1186/s40537-019-0197-0.
[17] X. Ying, "An overview of overfitting and its solutions," *Journal of Physics*, 2019.
[18] O. Ronneberger, P. Fischer, and T. Brox, *U-net: Convolutional networks for biomedical image segmentation*, 2015. arXiv: 1505.04597 [cs.CV].
[19] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
[20] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "Unet++: A nested u-net architecture for medical image segmentation," *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pp. 3–11, 2018.