Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Automated Histological Assessment of Parkinson's Disease Models via Integrated Dopaminergic Neuron Segmentation and Phenotypic Quantification

Prema Arokia Mary G
RV Institute of Technology and Management
Jayavadivel Ravi
infotech.jaya@gmail.com

Presidency University
Hema M S
RV Institute of Technology and Management
M Nageswara Guptha
Ramaiah University of Applied Science

Research Article

Keywords: Biochemical profiling, Dopaminergic neurons, Histopathology, Parkinson's disease, Semantic segmentation, Transformer architecture

Posted Date: July 15th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10278708/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Automated Histological Assessment of Parkinson's Disease Models via Integrated Dopaminergic Neuron Segmentation and Phenotypic Quantification

G Prema Arokia Mary[1], Hema M S[2], Jayavadivel Ravi[3], M Nageswara Guptha[4]

[1][2]Department of Computer Science and Engineering,

RV Institute of Technology and Management,

Bengaluru, Karnataka, India.

[3]Presidency University, Bengaluru, Karnataka, India

[4] Ramaiah University of Applied Science, Bengaluru, Karnataka, India.

Email: [1] premaamg.rvitm@rvei.edu.in,[2]ghema_shri@yahoo.co.in,

[3]infotech.jaya@gmail.com, [4]nageswaraguptha.cs.et@msruas.ac.in

Abstract

Accurate quantification of dopaminergic neurons and their phenotypic state is essential for understanding neurodegeneration and evaluating therapeutic efficacy in Parkinson's disease models. Traditional approaches, such as stereological counting, are labor-intensive and limited to structural assessment, while most automated methods neglect key biochemical markers like Tyrosine Hydroxylase (TH), a critical indicator of dopaminergic identity and viability. To overcome the limitations of existing methods that either focus solely on morphological segmentation or lack integrated biochemical analysis, a novel dual-head transformer-based architecture, SegDAFormer, is introduced in this research. It is specifically designed to jointly perform accurate dopaminergic neuron segmentation and fine-grained, pixel-wise quantification of TH expression from high-resolution histological images, thereby unifying structural and phenotypic assessment within a single, biologically interpretable framework. The model integrates a Hierarchical Transformer Encoder with a Multi-Scale Integrative Module, enabling effective fusion of local and global contextual features. Evaluated on a real TH-stained dataset, SegDAFormer achieves a Dice score of $93.12 \pm 1.08\%$, counting error of $5.82 \pm 1.45\%$, and TH intensity prediction error of MAE 7.96 ± 1.22, significantly outperforming existing baselines. Moreover, visual overlays confirm that predicted intensity maps reflect biologically meaningful gradients in neuronal phenotype. By bridging structural and biochemical analysis in a unified, interpretable framework, SegDAFormer advances automated histological assessment and holds strong potential for scalable applications in neurodegenerative research, drug screening, and digital pathology.

Key words: Biochemical profiling, Dopaminergic neurons, Histopathology, Parkinson's disease, Semantic segmentation, Transformer architecture

1. INTRODUCTION

Parkinson's disease manifests as motor dysfunctions such as tremor, rigidity, and bradykinesia, driven by progressive degeneration of dopamine-producing neurons in the substantia nigra pars compacta [1]. An estimated 6.1 million individuals were living with Parkinson's disease worldwide in 2016, a figure expected to surpass 12 million by 2040 as populations age. The need for methodologies capable of accurately quantifying dopaminergic neuron loss has been heightened by the increasing global burden of PD. Tyrosine Hydroxylase (TH) has been employed as the principal biochemical marker for dopaminergic neuron visualization and quantification [2].

Given that TH consistently marks dopaminergic neurons in both normal and pathological samples, quantifying TH-positive cells facilitates preclinical assessments of disease progression and the efficacy of interventions [3]. It is essential in preclinical studies to evaluate disease progression and therapeutic efficacy. As TH labeling robustly delineates dopaminergic neurons across healthy and diseased tissues, TH-positive cell enumeration has been established as an indispensable metric for tracking disease progression and assessing therapeutic efficacy in preclinical studies [4]. In experimental models of Parkinson's disease, researchers commonly assess not just the presence of TH-positive neurons, but also the intensity and distribution of TH expression to assess neuronal health [5]. A drop in TH signal typically reflects cell loss or diminished dopaminergic function, that are essential for judging the impact of potential therapies [6].

As TH expression reflects both the presence and health of dopaminergic neurons, analytical methods used to measure these cells captures not just their shape and location, but also the quality of their staining. This becomes especially important in larger studies, where researchers need to analyze high-resolution histological images quickly and consistently [7]. As these studies scale up, relying on manual methods becomes less practical. Consequently, there is an increasing demand for automated platforms capable of simultaneously characterizing neuronal architecture and biochemical markers, while preserving the fine-grained biological detail essential for in-depth analysis.

Conventional approaches, including manual stereological counting, though biologically valid, are labor-intensive, slow, and prone to subjectivity [8]. These limitations have created a growing demand for automated, scalable, and reproducible methods that can reliably process high-resolution histological data. However, automation of this task remains challenging due to complex neuronal morphologies, overlapping structures, and variability in staining intensity [9]. Deep learning models, such as U-Net [10] , Mask R-CNN [11], and Cellpose [12] have addressed these challenges by automating neuron detection and segmentation, thus facilitating more precise neuron quantification.

Recent advances in self-supervised learning have demonstrated that robust features can be extracted from unlabeled histology images, thereby reducing reliance on costly manual annotations and

yielding representations that generalize across datasets [13–14]. However, large annotated datasets are still required by most CNN-based approaches to dopaminergic neuron analysis. Further, as these models operate within confined spatial contexts, overlapping cellular structures are frequently misinterpreted. Essential phenotypic signals—such as variations in TH staining intensity—are also often neglected, compromising evaluations of neuronal health. Although throughput has been increased through the adoption of whole-slide processing and self-supervision, cohesive frameworks that integrate structural segmentation with biochemical quantification are not yet available. Consequently, the development of scalable, reproducible, and biologically transparent tools for Parkinson's disease preclinical research has been hindered.

In response to these shortcomings, SegDAFormer has been introduced as an end-to-end Transformer-based framework for the simultaneous segmentation of TH-positive neurons and prediction of their staining intensity. Prior convolutional methods have been hindered by several critical gaps: reliance on vast manually annotated datasets that limit scalability; restricted receptive fields that cause boundary errors in densely packed neuron clusters; compartmentalized pipelines that separate morphology and intensity analyses, thereby propagating errors and increasing processing time; absence of multi-scale feature integration, which prevents detection of subtle phenotypic gradients in staining; and lack of built-in explainability, which obscures decision rationale and hinders biological interpretation. SegDAFormer specifically addresses each of these gaps through a dual-output architecture with an integrated multi-scale feature fusion module, adaptive task weighting, and embedded explainability mechanisms. As a result, more precise, transparent, and biologically meaningful profiles of dopaminergic neuron populations are generated, directly overcoming the limitations of existing approaches.

The key contributions of this research are as below.

- A novel architecture to simultaneously segment dopaminergic neurons and quantify TH intensity, enabling integrated structural and phenotypic analysis within a single framework.
- A new fusion module that combines representations across multiple scales, ensuring the network retains fine-grained morphological features as well as the broader contextual cues needed for precise analysis.
- A novel joint optimization strategy to balance segmentation and phenotypic regression objectives, facilitating end-to-end learning that aligns with biological interpretation and experimental relevance.

To validate these contributions, SegDAFormer was evaluated on publicly available TH-stained histology images. Architectural ablations were performed to assess the impact of each model component, and cross-validation was used to confirm consistent performance across varied staining conditions and

neuronal densities. Explainability analyses were then employed to verify that both segmentation and profiling outputs aligned with known biological patterns.

The remainder of the paper is structured as follows: in Section 2, a comprehensive review of literature relevant to automated segmentation and phenotypic analysis of dopaminergic neurons is provided. The dataset employed in this study, together with details of implementation of the proposed model are described in Section 3. The proposed model architecture and the mathematical formulation of the joint optimization problem are detailed in Section 4. The results of empirical evaluations, including performance comparisons with existing methods, are reported in Section 5; and the paper is concluded in Section 6, where potential avenues for future research are also outlined.

2. RELATED WORKS

Substantial evolution has been observed in the automated analysis of TH-positive dopaminergic neurons in Parkinson's disease models, with workflows progressing from manual, design-based stereology to advanced deeplearning segmentation techniques. In this section, key developments in traditional counting approaches, CNN and Transformerbased segmentation frameworks, and emerging phenotypic profiling methods are reviewed, and the resulting limitations that motivated the proposed SegDAFormer are highlighted.

For decades, neuroscientists have relied on meticulous, design-based stereology to count dopaminergic (TH-positive) neurons, a method that is still regarded as the gold standard for its statistical validity and unbiased accuracy. Techniques such as the optical fractionator and physical disector allow researchers to systematically sample brain tissue and derive reliable neuron counts across defined anatomical regions. It was demonstrated by Ip et al. [15] that even a basic light microscope fitted with a motorized stage produces outcomes on par with those from high-end stereology platforms, thereby emphasizing the technique's accessibility and robustness. However, complete brain analysis can require days or weeks, as tissue preparation, sample selection, and manual counting are both time- and labor-intensive.

Moreover, no information on neuronal phenotype is furnished by traditional stereology, which reports only cell counts and typically omits markers such as TH staining intensity [16]. Although efficiency has been enhanced through automated sampling systems, digital slide scanners [17], and light-sheet microscopy [18], these methods remain focused on structural data, leaving neuronal physiological states largely unexamined. The laborious nature of these workflows, their susceptibility to human error, and their limited throughput render them impractical for large-scale, high-resolution studies in neurodegeneration research. To address these shortcomings, deep learning approaches have been introduced, offering automated, scalable, and more refined neuron segmentation solutions [19–20].

Building on the U-Net foundation, several research groups have further customized CNN pipelines for dopaminergic neuron detection and analysis [21].

For instance, Barzekar et al. [22] introduced a U-Net–based multiclass semantic segmentation pipeline that employed a ResNet-34 encoder and extensive augmentation strategies to delineate anatomical sub-regions of the substantia nigra and infer neuronal health, achieving a mean Intersection over Union (IoU) of 79 % and a mean Dice coefficient of 87 %. Building on CNN architectures, Kim et al. [23] implemented a custom convolutional framework tailored for TH-positive dopaminergic neuron detection in mouse models; their model incorporated batch normalization and dropout layers to improve robustness under variable staining conditions, and demonstrated classification accuracy outperforming traditional thresholding methods (reported accuracies > 90 %). Meanwhile, Pandey et al. [24] leveraged transfer learning using pre-trained EfficientNet-B0 backbones to predict Parkinson's disease progression from diverse data sources, including spiral drawings, achieving a top-1 diagnostic accuracy of approximately 92 %; however, their work focused on image-level classification instead of cellular segmentation. In a complementary direction, Chen et al. [25] developed TrueTH, a U-Net-derived architecture with integrated attention gates and a dual-branch output designed for simultaneous detection and phenotypic scoring of dopaminergic neurons; the model achieved high precision and recall (> 88 %) on a proprietary TH-stained dataset.

In a recent study, Wang et al. [26] utilized whole-brain 3D tissue clearing using Stabilization under Harsh Conditions via Intramolecular Epoxide Linkages (SHIELD) to prevent degradation and Clear, Unobstructed Brain Imaging Cocktails and Computational analysis (CUBIC) protocols, combined with volumetric confocal microscopy, to map TH-labeled dopaminergic neurons, glial cells, and vasculature in a 6-hydroxydopamine (6-OHDA) mouse model of Parkinson's disease. Their pipeline enabled region-specific quantification of TH signal loss and glial proliferation in areas such as the substantia nigra, ventral tegmental area, and caudate-putamen (e.g., reduced TH fluorescence and increased astrocyte density; $p < 0.05$), offering rich mesoscale phenotypic insights. However, it lacked an automated deep learning based segmentation framework for cell-level analysis. In this context, Fred et al. [27] applied a three-dimensional U-Net architecture to ultra-high-field 7 Tesla magnetic resonance imaging (7T MRI) scans for precise segmentation of the substantia nigra, aiming to enhance diagnostic accuracy in Parkinson's disease. While effective at anatomical localization using high-resolution imaging, their approach did not target cellular-scale features or phenotypic profiling, highlighting the need for integrating structural segmentation with molecular phenotype quantification at the single-cell level.

In addition to MRI-based anatomical segmentation, recent works have explored deep learning innovations with greater contextual alignment to cellular and histological applications. Basukala et al. [28] utilized Quantitative Susceptibility Mapping (QSM) with patch-wise classification and voxel-wise

refinement to segment the substantia nigra and red nucleus in Parkinson's disease patients, achieving 89% accuracy. While precise, their method remained restricted to MRI-level resolution without phenotypic granularity. Nebel et al. [29] introduced Conv-SeTr, a hybrid convolution-Transformer model that segmented midbrain nuclei on T2-weighted MRI with a Dice score of 0.81, outperforming conventional U-Nets, but again limited to macrostructural imaging.

In this line, Haghighi et al. [30] proposed a self-supervised U-Net framework trained on unlabeled TH-stained images to segment and quantify dopaminergic neurons. Their model significantly reduced annotation dependence and achieved Dice scores exceeding 85%, though phenotypic regression was handled post hoc. Wang et al. [31] employed Swin-Unet for substantia nigra segmentation in SWI and T2-weighted MRI, reporting a Dice score of 0.832 and an AUC of 0.894 for PD classification. Similarly, Hussain et al. [32] developed a Swin Transformer-CNN fusion network that reached 96% classification accuracy, highlighting the potential of hybrid attention models. Yet, none of these approaches jointly model neuron morphology and phenotypic intensity within a unified, interpretable architecture.

Recent advances in digital pathology are shifting focus from merely locating cells to uncovering how they behave and function at a deeper level. Lu et al. [33] introduced ERnet, a segmentation tool that maps the complex topology of the endoplasmic reticulum, showing how deep learning can quantify subtle phenotypic traits like branching and continuity. While not neuron-specific, it illustrates how structure can reflect cellular function. In a closer context, Di Credico et al. [34] used machine learning to detect morphological and phenotypic shifts in human dopaminergic neurons exposed to environmental chemicals, combining high-content imaging with clustering algorithms to identify subtle toxicity-induced changes.

On the disease modeling front, D'Sa et al. [35] used patient-derived stem cells and deep phenotyping to predict Parkinson's subtypes, linking transcriptomic profiles with cell morphology. Weiss et al. [36] took this further by systematically quantifying TH intensity and neurite outgrowth in stem cell-derived dopaminergic neurons, highlighting the need for scalable, reproducible phenotypic readouts. Villaronga-Luque et al. [37] found that changes in cell metabolism could directly influence how cells grow and organize themselves in stem-cell-based embryo models—showing just how closely cell shape and molecular activity are connected. Together, these studies reflect a shared goal in the field—to understand not just what cells look like but what their appearance reveals about their function.

Despite notable progress across segmentation and phenotypic profiling, a critical gap persists: no current framework offers a unified solution that simultaneously segments dopaminergic neurons and quantifies their biochemical state at the cellular level. Existing models typically specialize in either structural delineation or post hoc phenotype analysis, often relying on separate pipelines, handcrafted

features, or modality-specific constraints. This fragmentation limits both the interpretability and scalability of such tools in high-throughput neurodegenerative research.

There is a pressing need for an end-to-end approach that integrates multi-scale feature learning with dual-task optimization—capable of preserving cellular morphology, capturing phenotypic gradients like TH intensity, and generalizing across real-world histological variability. This research directly addresses these gaps by introducing a unified framework that couples Transformer-based segmentation with phenotypic profiling, enabling comprehensive, interpretable analysis of dopaminergic neurons at both structural and biochemical levels.

3. MATERIALS AND METHODS

In this section, the methodology underlying the proposed framework is described. The problem is first formally defined to establish the computational tasks of segmentation and phenotypic profiling. This is followed by a detailed description of the dataset used for training and evaluation, along with the implementation strategies adopted.

3.1 Problem Definition

This research formalizes dopaminergic neuron analysis as a dual-objective learning problem, where the model is trained to simultaneously predict a binary segmentation mask and a continuous TH intensity map for each input image. By embedding both structural and phenotypic tasks into a single optimization framework, the approach enables end-to-end learning of a function f_θ that captures not only the spatial extent of TH-positive neurons but also their biochemical state at the pixel level. This formulation serves as the foundation for the mathematical definition that follows.

Let $\mathcal{X} = \{x_i\}_{i=1}^{N}$ denote a set of high-resolution histological images where each $x_i \in \mathbb{R}^{H\times W\times 3}$ represents an RGB image of tissue stained for TH. The goal is to learn a function f_θ parameterized by θ , which jointly performs two tasks:

1. **Binary segmentation**: Predict a segmentation mask $\hat{y}_i^{(seg)} \in \{0,1\}^{H\times W}$ for each input x_i , where pixels labeled as 1 correspond to dopaminergic neuron somata.
2. **Phenotypic profiling**: Predict a continuous-valued intensity map $\hat{y}_i^{(int)} \in [0,1]^{H\times W}$ reflecting relative TH expression at each pixel.

Given corresponding ground truth masks $\hat{y}_i^{(seg)}$ and intensity maps $y_i^{(int)}$, the model is trained to minimize a joint loss function as in (1), where $\mathcal{L}_{\text{seg}}$ is a segmentation loss combining Dice and Focal loss, $\mathcal{L}_{\text{int}}$is the Smooth L_1 (Huber) loss for regression,λ_1 and λ_2 are weighting coefficients balancing the two objectives.The model is optimized end-to-end to learn shared and task-specific representations that enable accurate morphological segmentation and biologically meaningful intensity quantification of dopaminergic neurons.

$$\mathcal{L}_{\text{total}} = \lambda_1 \cdot \mathcal{L}_{\text{seg}}(y_i^{(\text{seg})}, \hat{y}_i^{(\text{seg})}) + \lambda_2 \cdot \mathcal{L}_{\text{int}}(y_i^{(\text{int})}, \hat{y}_i^{(\text{int})}) \quad (1)$$

3.2 Dataset Description

The dataset utilized in this study is the first publicly available collection specifically curated for the segmentation and quantification of TH-positive dopaminergic neurons in mouse brain tissue [38]. It comprises 114 high-resolution digital histology images acquired from mouse brains sectioned at 35 µm thickness and immunostained for TH, a key enzyme marker of DA neurons. Either Haematoxylin or Nissl stains were applied as counterstains to enhance tissue structure visibility. Imaging was performed using a Hamamatsu Nanozoomer whole-slide scanner at 20x magnification, achieving a spatial resolution of 0.46 µm/pixel.Each image in the dataset includes an associated 16-bit labeled mask, where individual DA neuron somas are annotated. In total this dataset has approximately 18,193 DA neuron cell bodies.

The dataset is split equally into training and test subsets, each comprising 57 images. To amplify morphological variability, the training subset is enriched by applying ten discrete augmentations to each image: four right-angle rotations (0°, 90°, 180°, 270°), a horizontal flip (p = 1.0), a vertical flip (p = 1.0), a random resized crop (scale = 0.8–1.0, aspect ratio = 3/4–4/3), an RGB channel shift (±10 intensity units), a Gaussian blur (σ = 0.1–2.0 px), and Gaussian noise (σ = 0–0.05 of the intensity range). A total of 570 augmented images is thus generated, resulting in a composite training pool of 627 samples, which is subsequently partitioned into model training (80%, n = 502) and internal validation (20%, n = 125) subsets. The test subset remains unaltered to safeguard the integrity of performance evaluation.Table 1gives the augmentation summary and the distribution of the dataset.

Table1. Dataset Split and Augmentation Summary

Subset	# Originals	Augmentation (per image)	# Augmented	Total Samples	Split Ratio
Training Pool	57	Rotation (0°, 90°, 180°, 270°) H-flip (p=1.0) V-flip (p=1.0) ResizedCrop (0.8–1.0, 3/4–4/3) RGBShift (±10) Blur (σ 0.1–2.0 px) Noise (σ 0–0.05)	570	627	100% of training pool
Model Training	—	—	—	502	80% of training pool
Internal Validation	—	—	—	125	20% of training pool
Test	57	None	0	57	50% of dataset
Total	114	—	570	684	100% of dataset

3.3 Details of Implementation

All computational experiments and analyses in this study were performed using MATLAB R2024b. The platform provided a robust environment for image processing, post-segmentation analysis, and performance evaluation. Key MATLAB toolboxes used include the Image Processing Toolbox for handling high-resolution histology images and the Parallel Computing Toolbox to leverage GPU acceleration and parallelize data-intensive operations.

The implementation was executed on a high-performance workstation running Windows 10 Pro (64-bit). The system was equipped with an Intel® Xeon® W-2295 CPU operating at 3.00 GHz with 18 cores, paired with 256 GB of DDR4 RAM to efficiently manage large volumetric data and memory-intensive computations. For GPU-accelerated tasks, an NVIDIA RTX A6000 graphics card with 48 GB VRAM was utilized. Storage requirements for whole-slide image handling and rapid I/O were met using a 2 TB NVMe SSD, ensuring high-speed data throughput.

4. SegDAFormer ARCHITECTURE

This section presents the design and functioning of SegDAFormer, a unified dual-headed architecture tailored for the joint segmentation and phenotypic profiling of dopaminergic (DA) neurons. Built on a modified SegFormer backbone, SegDAFormer integrates morphological and intensity-based representations through a coordinated multi-scale framework, aiming to deliver biologically meaningful outputs with high precision and interpretability. The model is optimized to handle challenges such as sparse staining, cell overlap, and inter-animal histological variability.Figure 1 illustrates the highlevel architecture of the SegDAFormer.



Figure 1. SegDAFormer Architecture

4.1 Preprocessing

All input images were first converted to 32-bit floating-point format, and each RGB channel was independently scaled to [0, 1] by subtracting its minimum and dividing by its dynamic range 1. To suppress high-frequency noise and staining artifacts, a Gaussian smoothing filter ($\sigma = 1.0$) followed by non-local means denoising was applied to each image. Each denoised image was then resized to 512 × 512 pixels via bicubic interpolation to standardize input dimensions. During training, on-the-fly augmentations were performed, including random horizontal and vertical flips, rotations, brightness/contrast jitter, elastic deformations, and additive Gaussian noise. Macenko stain normalization

was applied across the dataset to harmonize color distributions. Finally, each augmented image was normalized by subtracting the global mean and dividing by the global standard deviation. This pipeline ensured that inputs were denoised, uniformly scaled, and diverse, promoting stable model convergence and robust generalization.

4.2 Hierarchical Transformer Encoder and Feature Representation

The Hierarchical Transformer Encoder forms the foundational backbone of SegDAFormer, enabling the extraction of rich spatial-semantic features from high-resolution histological images. Unlike traditional convolution-based encoders, this module is designed to preserve both local structural detail and long-range contextual information through a multi-stage transformer framework inspired by theMix Transformer (MiT) architecture as shown in Figure 2.The following subsections present the internal structure, mathematical operations, and multi-scale feature representations produced by each stage of the Hierarchical Transformer Encoder.

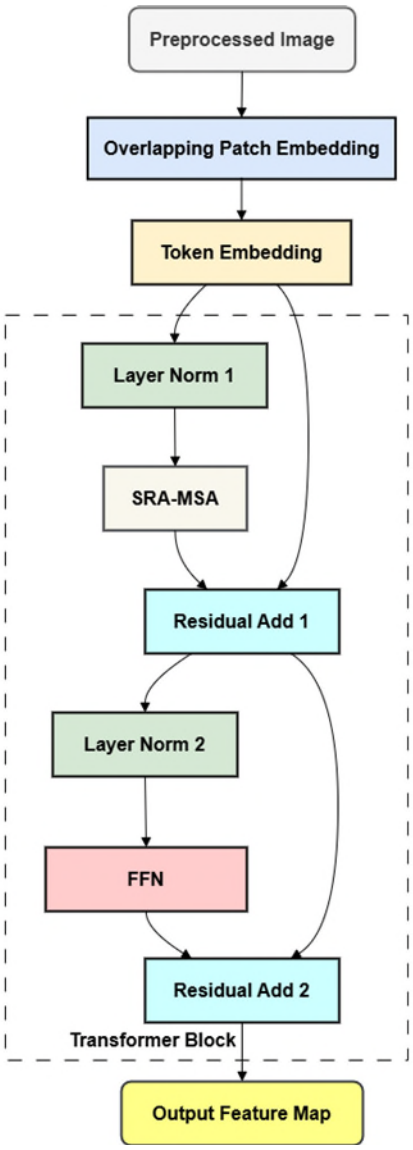


Figure 2. Hierarchical Transformer Encoder

a. Patch Embedding

Given a preprocessed input image $\mathbf{X} \in \mathbb{R}^{H \times W \times 3}$, the encoder begins with an overlapping patch embedding operation. This is implemented using a 2D convolutional layer with kernel size $k = 7$, stride $s = 4$, and padding $p = 3$, which transforms the image into a lower-resolution feature map as in (2). This produces an initial token representation while retaining spatial continuity via overlapping receptive fields. The output $\mathbf{E}_0$ is flattened and rearranged into a token sequence for transformer processing.

$$\mathbf{E}_0 = \text{Conv2D}(\mathbf{X}; k = 7, s = 4, p = 3) \text{where } \mathbf{E}_0 \in \mathbb{R}^{H' \times W' \times C} \quad (2)$$

b. Hierarchical Multi-Stage Encoding

The encoder comprises four parallel stages: P_1, P_2, P_3, P_4, each responsible for producing feature maps at decreasing spatial resolutions and increasing channel depths:

- P_1: resolution $H/4 \times W/4$, channels 128
- P_2: resolution $H/8 \times W/8$, channels 128
- P_3: resolution $H/16 \times W/16$, channels 320
- P_4: resolution $H/32 \times W/32$, channels 512

Each stage consists of a Transformer Block, detailed below. Each Transformer Block processes a sequence of tokens $\mathrm{Z} \in \mathbb{R}^{N \times D}$ where $N = H \times W$ and D is the feature dimension. It follows a two-branch structure, both involving Layer Normalization, residual connections, and either attention or feed-forward operations:

1. **First Sub-layer: Spatial Reduction Multi-Head Self Attention (SRA-MSA):** The first sub-layer applies layer normalization, followed by SRA-MSA. Unlike standard attention, SRA reduces the spatial dimensionality of key and value tokens via strided convolution, reducing computational load as in (3).

$$\mathbf{Z}' = \text{SRA-MSA}\big(\text{LN}_1(\mathbf{Z})\big) \text{ and } \mathbf{Z}_1 = \mathbf{Z} + \mathbf{Z}' \quad (3)$$

Multi-head self-attention is defined as in (4), where

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^{\top}}{\sqrt{d_k}}\right) \quad (4)$$

- $Q = \mathbf{Z}W_Q, K = \mathbf{Z}W_K, V = \mathbf{Z}W_V$
- W_Q, W_K, W_V are learnable projection matrices

In SRA, K and V are computed on spatially downsampled tokens, reducing complexity from $\mathcal{O}(N^2)$ to approximately $\mathcal{O}(N \cdot \sqrt{N})$.

2. **Second Sub-layer: Feed-Forward Network (FFN):** The second sub-layer applies a position-wise two-layer MLP with Gaussian Error Linear Unit(GELU) activation, followed again by residual addition as in (5). The feed-forward network is defined as in (6), where W_1 and W_2 are learned linear projections, and GELU is the Gaussian Error Linear Unit activation.

$$\mathbf{Z}'' = \text{FFN}(\text{LN}_2(\mathbf{Z}_1)) \text{and} \mathbf{Z}_{\text{out}} = \mathbf{Z}_1 + \mathbf{Z} \quad (5)$$

$$\text{FFN}(x) = W_2 \cdot \text{GELU}(W_1 x + b_1) + b_2 \quad (6)$$

3. **Output of the Encoder:** Each stage outputs a feature map $\mathbf{F}_i$ where $i \in \{1,2,3,4\}$, such that $\mathbf{F}_1 \in \mathbb{R}^{H/4\times W/4\times 128}$, $\mathbf{F}_2 \in \mathbb{R}^{H/8\times W/8\times 128}$, $\mathbf{F}_3 \in \mathbb{R}^{H/16\times W/16\times 320}$ and $\mathbf{F}_4 \in \mathbb{R}^{H/32\times W/32\times 512}$ representing spatial hierarchies at different resolutions as below. These multi-resolution feature maps are forwarded to the Multi-Scale Integrative Module (MSIM), which performs resolution alignment and feature fusion for downstream tasks such as segmentation and phenotypic profiling.

4.3 Multi-Scale Integrative Module

This module is designed to unify heterogeneous feature representations extracted from multiple stages of the Hierarchical Transformer Encoder. This fusion enables SegDAFormer to effectively capture both local fine-grained details and high-level contextual patterns, essential for segmenting morphologically diverse and overlapping dopaminergic neurons. The schematic of the MSIM is illustrated with Figure 3 and the steps outlined below describe the internal operations of the MSIM, which transforms multi-resolution encoder outputs into a unified feature representation for downstream decoding.

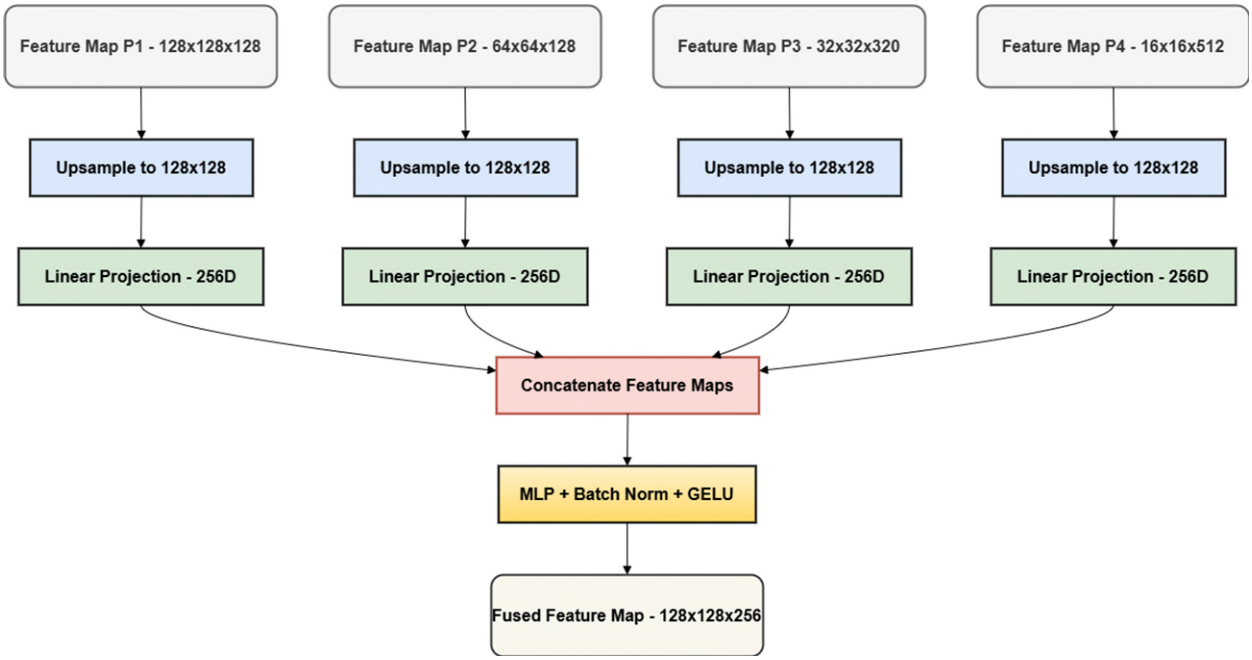


Figure 3. Multi-Scale Integrative Module

Let the input feature maps from stages P1 to P4 be denoted as: $\mathbf{F}_1 \in \mathbb{R}^{128\times 128\times 128}$, $\mathbf{F}_2 \in \mathbb{R}^{64\times 64\times 128}$, $\mathbf{F}_3 \in \mathbb{R}^{32\times 32\times 320}$ and $\mathbf{F}_4 \in \mathbb{R}^{16\times 16\times 512}$.

Step 1: Spatial Alignment via Upsampling: Each $\mathbf{F}_i$ is resized to a common spatial resolution of 128×128 using bilinear interpolation as in (7).

$$\hat{\mathbf{F}}_i = \text{Upsample}(\mathbf{F}_i, \text{size} = 128 \times 128) \text{for } i = 1,2,3,4 \quad (7)$$

Step 2: Channel Projection via Linear Transformation: Each upsampled feature map is passed through a 1×1 convolution (linear projection) to unify the channel dimensionality to $d = 256$, as in (8), which yields $\tilde{F}_i \in \mathbb{R}^{128\times128\times256}$.

$$\tilde{\mathbf{F}}_i = \hat{\mathbf{F}}_i \cdot \mathbf{W}_i + \mathbf{b}_i x \quad \text{where } \mathbf{W}_i \in \mathbb{R}^{C_i\times256} \tag{8}$$

Step 3: Feature Concatenation: The projected features are concatenated across the channel dimension as in (9).

$$\mathbf{F}_{\text{concat}} = \text{Concat}(\tilde{\mathbf{F}}_1, \tilde{\mathbf{F}}_2, \tilde{\mathbf{F}}_3, \tilde{\mathbf{F}}_4) \in \mathbb{R}^{128\times128\times(4\times256)} = \mathbb{R}^{128\times128\times1024} \tag{9}$$

Step 4: Nonlinear Fusion:The concatenated tensor is passed through a Multilayer Perceptron (MLP) with Batch Normalization and GELU activation to produce the fused feature map as in (10), which generates the final output $F_{\text{fused}} \in \mathbb{R}^{128\times128\times256}$.

$$\mathbf{F}_{\text{fused}} = \text{GELU}(\text{BN}(\mathbf{F}_{\text{concat}} \cdot \mathbf{W}_{\text{fuse}} + \mathbf{b}_{\text{fuse}})) \quad \text{where } \mathbf{W}_{\text{fuse}} \in \mathbb{R}^{1024\times256} \tag{10}$$

4.4 Segmentation Head

The Segmentation Head in SegDAFormer is designed to predict binary masks of dopaminergic neuron somas from the fused feature representation produced by the MSIM. Its design combines spatial refinement, boundary awareness, and multi-scale context aggregation to robustly identify dopaminergic neuron somas, even in the presence of morphological overlap or weak staining. The following steps describe the internal flow of the Segmentation Head, and Figure 4 depicts its architectural layout.

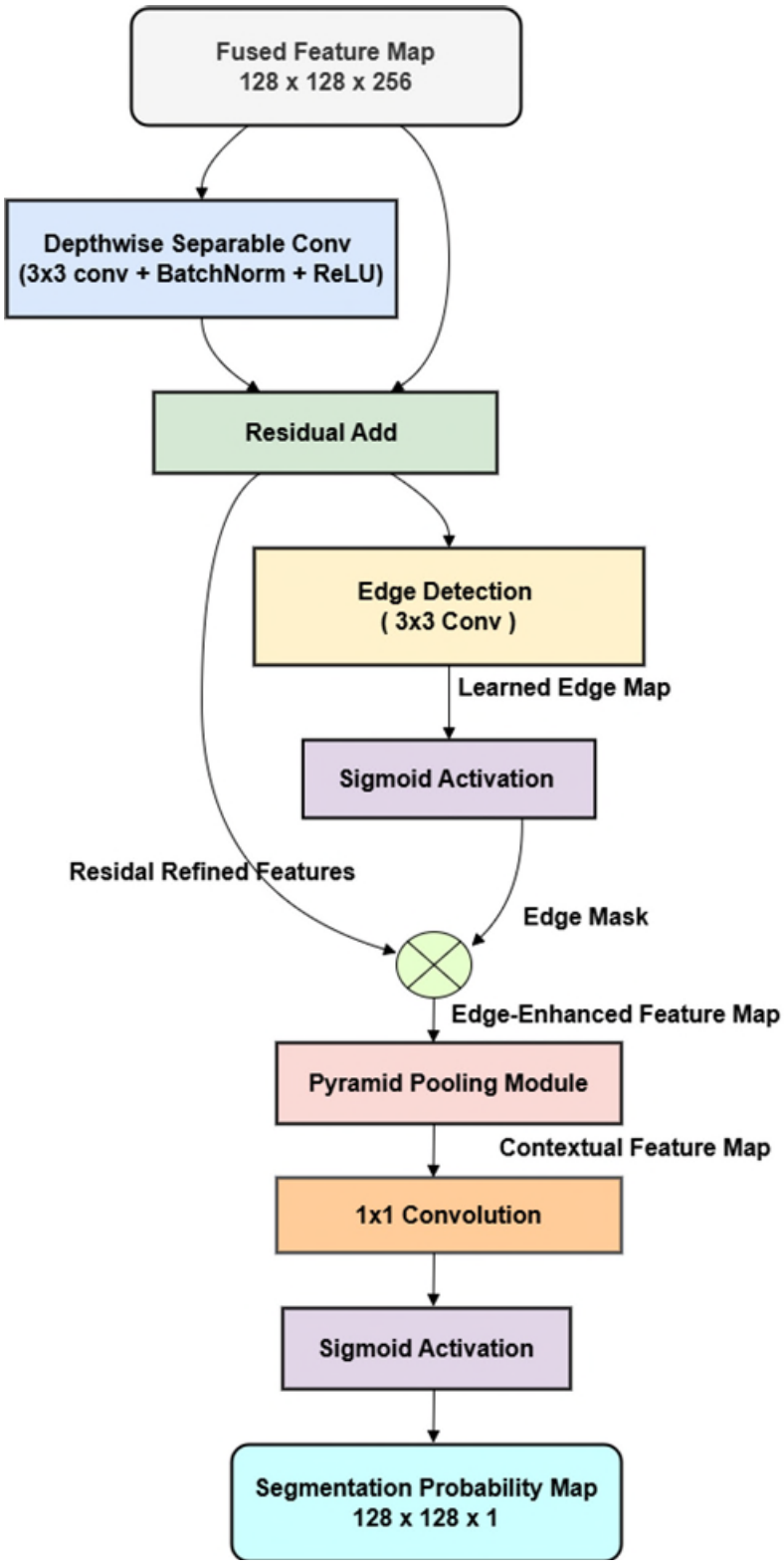


Figure 4. Segmentation Head

Step 1: Depthwise Separable Convolution: A depthwise separable convolution followed by Batch Normalization and ReLU is used to enhance spatial resolution of $\mathbf{F}_{\text{fused}}$ while maintaining computational efficiency as in (11).

$$\mathbf{F}_r = \text{ReLU}(\text{BN}(\text{DWConv}_{3\times3}(\mathbf{F}_{\text{fused}}))) \tag{11}$$

Step 2: Residual Connection : The refined output is summed with the original fused map to preserve global context as in (12).

$$\mathbf{F}_{\text{res}} = \mathbf{F}_r + \mathbf{F}_{\text{fused}} \tag{12}$$

Step 3: Edge Detection and Attention: A 3×3 convolution is applied to $\mathbf{F}_{\text{res}}$ to extract boundary information as in (13). This produces a soft edge mask with values in [0,1] indicating edge saliency. The refined feature is modulated by the edge mask as in (14).

$$\mathbf{E}_{\text{mask}} = \sigma(\text{Conv}_{3\times3}(\mathbf{F}_{\text{res}})) \tag{13}$$

$$\mathbf{F}_{\text{edge}} = \mathbf{F}_{\text{res}} \odot (1 + \mathbf{E}_{\text{mask}}) \tag{14}$$

Step 4: Pyramid Pooling for Contextual Awareness: The edge-weighted feature map is passed through a Pyramid Pooling Module (PPM) that pools features at multiple scales (e.g., 1x1, 2x2, 3x3, 6x6), upsamples them to 128×128, and concatenates them to enhance global context. The output of PPM is as in (15).

$$\text{F}_{\text{ppm}} \in \mathbb{R}^{128\times128\times256} \tag{15}$$

Step 5: Final Projection and Sigmoid: A 1×1 convolution reduces the feature map to a single channel as in (16). Then a sigmoid activation converts the logits into a probabilistic mask as in (17). The final binary segmentation mask $\mathbf{S}_{\text{bin}}$ is constructed by applying a threshold τ to the predicted probability map $\mathbf{S}_{\text{pred}}$, such that $\mathbf{S}_{\text{bin}} = \mathbb{I}(\mathbf{S}_{\text{pred}} \geq \tau)$, where $\mathbb{I}(\cdot)$denotes the indicator function.

$$\mathbf{S}_{\text{logit}} = \text{Conv}_{1\times1}(\mathbf{F}_{\text{ppm}}) \tag{16}$$

$$\mathbf{S}_{\text{pred}} = \sigma(\mathbf{S}_{\text{logit}}) \text{where} \mathbf{S}_{\text{pred}} \in \mathbb{R}^{128\times128\times1} \tag{17}$$

4.5 Phenotypic Profiling Head

The Phenotypic Profiling Head is designed to estimate TH intensity as a proxy for dopaminergic neuron phenotype, complementing the segmentation output by quantifying neuronal viability. As illustrated in Figure 5, the architecture comprises two primary branches: one produces a high-resolution pixel-wise TH intensity map to capture local phenotypic variation, while the other performs masked, multi-scale pooling to derive scalar estimates of local, regional, and global expression. A parallel confidence estimation branch quantifies the reliability of these predictions. The following sections describe each component in detail.

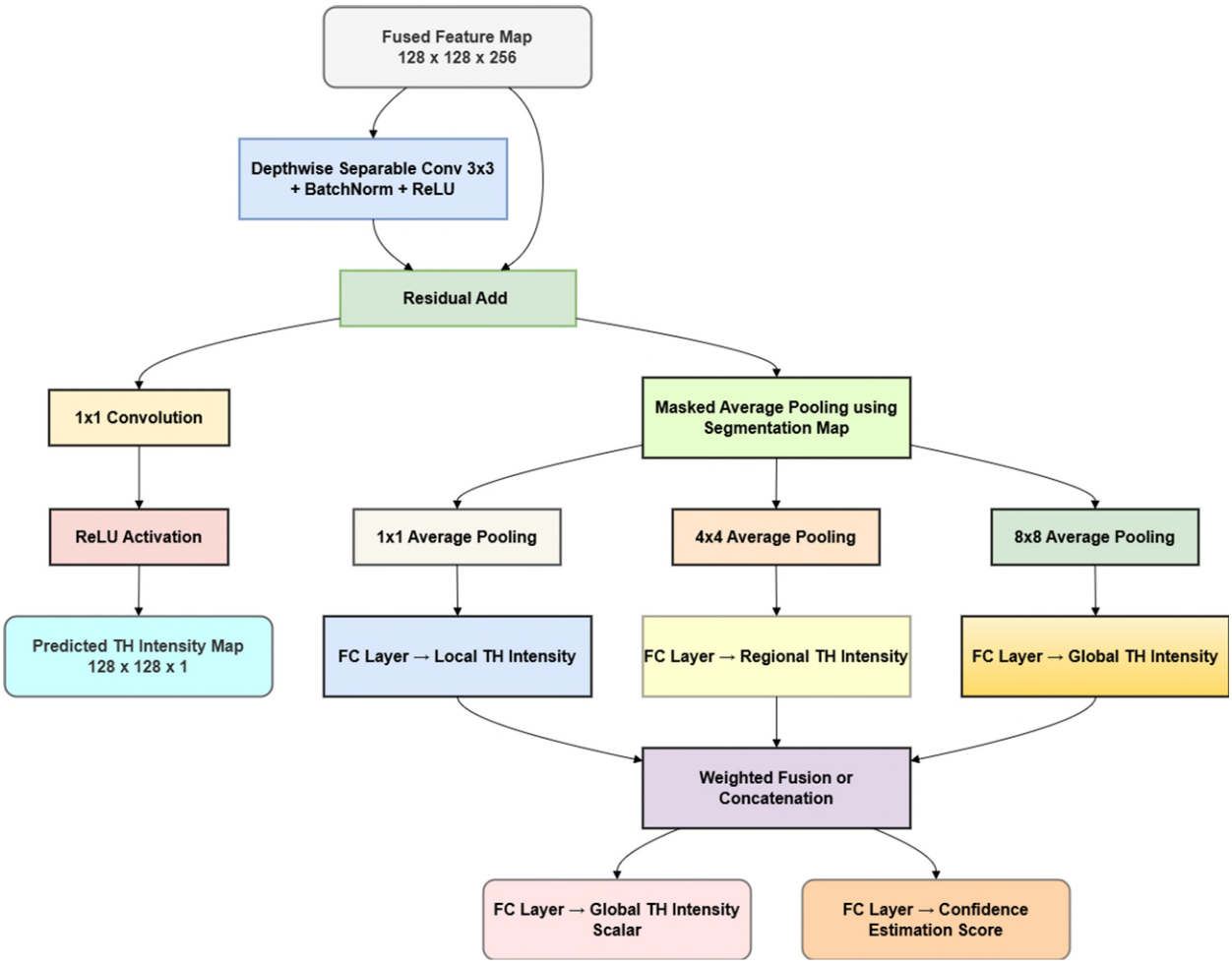


Figure 5. Phenotypic Profiling Head

The input to the profiling head is the fused feature map $\mathbf{F}_{\text{fused}}$ from the MSIM. A depthwise separable convolution, followed by Batch Normalization and ReLU activation, refines the input while preserving spatial structure as in (18). A residual connection is applied to integrate the refined and original features as in (19).

$$\mathbf{F}_{\text{ref}} = \text{ReLU}(\text{BN}(\text{DWConv}_{3\times3}(\mathbf{F}_{\text{fused}}))) \tag{18}$$

$$\mathbf{F}_{\text{sum}} = \mathbf{F}_{\text{ref}} + \mathbf{F}_{\text{fused}} \tag{19}$$

a. **Pixel-Wise TH Intensity Map:** To retain spatial resolution, a dedicated branch outputs a pixel-level intensity map. A 1×1 convolution followed by ReLU activation ensures non-negative predictions as in (20). This map serves as a direct estimate of pixel-wise TH intensity and can be used for spatial overlays, statistical summaries, or threshold-based phenotypic classification.

$$\mathbf{I}_{\text{map}} = \text{ReLU}(\text{Conv}_{1\times1}(\mathbf{F}_{\text{sum}})) \text{where } \mathbf{I}_{\text{map}} \in \mathbb{R}^{128\times128\times1} \tag{20}$$

b. **Multi-Scale Regional TH Estimation:** In parallel, a second branch performs masked average pooling, guided by the binary segmentation mask S_{bin}, to aggregate phenotypic features only from soma-

positive regions as in (21). To capture multi-scale context, three pooling operations are performed:$\text{AvgPool}_{1\times1}, \text{AvgPool}_{4\times4}, \text{AvgPool}_{8\times8}$.

$$\mathbf{F}_{\text{masked}} = \frac{1}{\sum_{i,j} \mathbf{S}_{\text{bin}}(i,j)} \sum_{i,j} \mathbf{F}_{\text{sum}}(i,j) \cdot \mathbf{S}_{\text{bin}} \tag{21}$$

Each output is passed through a Fully Connected (FC) layer to yield local, regional, and global intensity estimates as in (22-24).These predictions are concatenated and processed via another FC layer to yield the final global TH intensity value as in (25).

$$\mathbf{I}_{\text{local}} = \text{FC}_1(\text{AvgPool}_{1\times1}(\mathbf{F}_{\text{masked}})) \tag{22}$$

$$\mathbf{I}_{\text{regional}} = \text{FC}_2(\text{AvgPool}_{4\times4}(\mathbf{F}_{\text{masked}})) \tag{23}$$

$$\mathbf{I}_{\text{global}} = \text{FC}_3(\text{AvgPool}_{8\times8}(\mathbf{F}_{\text{masked}})) \tag{24}$$

$$\mathbf{I}_{\text{final}} = \text{FC}_{\text{fuse}}([\mathbf{I}_{\text{local}}, \mathbf{I}_{\text{regional}}, \mathbf{I}_{\text{global}}]) \tag{25}$$

4.6 Confidence Estimation

To quantify the model's reliability, a parallel confidence estimation branch predicts an uncertainty score associated with the final TH intensity. This scalar output is produced by an additional FC layer applied to the fused pooled features as in (26). This score serves as a soft measure of trust in the phenotypic readout and can aid downstream analysis in weakly stained or ambiguous regions.

$$\mathbf{U} = \text{FC}_{\text{conf}}([\mathbf{I}_{\text{local}}, \mathbf{I}_{\text{regional}}, \mathbf{I}_{\text{global}}]) \tag{26}$$

In summary, SegDAFormer introduces a unified architecture that jointly performs structural segmentation and phenotypic profiling of dopaminergic neurons from TH-stained histological images. Its hierarchical transformer encoder extracts multi-scale semantic features, which are harmonized through the MSIM for comprehensive spatial representation. The dual-head decoding structure enables precise morphological segmentation and biologically meaningful quantification of TH intensity, both locally and globally. By integrating pixel-level and region-level outputs, along with confidence estimation, SegDAFormer offers a robust and interpretable framework tailored for high-throughput and biologically relevant analysis in Parkinson's disease models.

4.7 Model Training

SegDAFormer is trained in a supervised, multi-task learning setting, jointly optimizing for structural segmentation and phenotypic profiling. The architecture is designed to learn both morphological boundaries and biochemical intensities, requiring distinct yet coordinated objectives. The training process employs task-specific loss functions and a composite optimization objective to ensure stable convergence and effective feature sharing between heads.

For the segmentation head, a compound loss function $\mathcal{L}_{\text{seg}}$ is employed, combining Dice Loss and Focal Loss. The Dice component promotes region-level accuracy, particularly beneficial for small or fragmented neuron masks, while the Focal component mitigates class imbalance by emphasizing hard-to-classify pixels. The total segmentation loss is defined as in (27), where λ_1 and λ_2 are learnable scalar weights that are optimized jointly with the model parameters. This allows the network to adaptively balance region overlap and pixel-level discrimination throughout training. The Dice Loss is expressed as in (28) and the and the Focal Loss is formulated as in (29), where $\hat{y}_i$and y_i are the predicted and ground truth binary labels, α is the class weighting factor, γ is the focusing parameter, and ϵ is a smoothing term.

$$\mathcal{L}_{\text{seg}} = \lambda_1 \cdot \mathcal{L}_{\text{Dice}} + \lambda_2 \cdot \mathcal{L}_{\text{Focal}} \tag{27}$$

$$\mathcal{L}_{\text{Dice}} = 1 - \frac{2\sum_i \hat{y}_i y_i + \epsilon}{\sum_i \hat{y}_i + \sum_i y_i + \epsilon} \tag{28}$$

$$\mathcal{L}_{\text{Focal}} = -\sum_i \alpha(1-\hat{y}_i)^{\gamma} y_i \log(\hat{y}_i) \tag{29}$$

The profiling head is trained using the Smooth L1 Loss (Huber loss), which is more robust to outliers than Mean Squared Error (MSE) and ensures stable regression of TH intensity values across varying expression levels. It is defined as in (30), where r_j and $\hat{r}_j$ are the ground truth and predicted TH intensities for pixel or region j, and M is the total number of intensity targets.

$$\mathcal{L}_{\text{reg}} = \frac{1}{M}\sum_{j=1}^{M} \text{SmoothL1}(r_j - \hat{r}_j) \tag{30}$$

where $\text{SmoothL1}(x) = \begin{cases} 0.5x^2, & \text{if } |x| < 1 \\ |x| - 0.5, & \text{otherwise} \end{cases}$

The final objective is a weighted sum of the segmentation and profiling losses as defined in (31). The model parameters θ are optimized by minimizing the total loss over the training set as in (32), where x is the input image, y is the segmentation mask, r is the TH intensity target, and α,β, are task-balancing weights.

$$\mathcal{L}_{\text{total}} = \alpha \cdot \mathcal{L}_{\text{seg}} + \beta \cdot \mathcal{L}_{\text{reg}}] \tag{31}$$

$$\theta^* = \arg\min_{\theta} \mathbb{E}_{x,y,r}[\mathcal{L}_{\text{total}}(x, y, r; \theta)] \tag{32}$$

This adaptive, multi-objective training strategy enables SegDAFormer to learn both structural and phenotypic representations in a balanced and biologically meaningful manner. Robust generalization across histological datasets exhibiting diverse neuronal morphologies was achieved by jointly optimizing the segmentation and phenotypic profiling loss functions.

5. EXPERIMENTS AND RESULTS AND DISCUSSIONS

To evaluate SegDAFormer under realistic conditions, a series of experiments was conducted to assess accuracy, robustness, and interpretability. Quantitative evaluations of neuron segmentation and TH

intensity estimation were performed, and performance was benchmarked against established models. Ablation studies were executed to isolate the effects of individual architectural components, while explainability analyses were carried out to highlight the model's decision pathways. Results are presented in the following sections, with a subsequent discussion of their biological significance and connections to current literature.

5.1. Experimental Setup

For optimal generalizability, SegDAFormer was subjected to grid-search hyperparameter tuning combined with five-fold cross-validation. Learning rates, task weights, and regularization strengths were varied across folds, each of which involved training on four data partitions and validating on the held-out partition. Mean Dice scores for segmentation and Smooth L1 loss for TH regression were averaged over all folds to guide the selection of hyperparameters. Finally, the relative weighting of segmentation and profiling objectives was adjusted to ensure balanced optimization during multi-task training.The tuning process prioritized segmentation quality (measured by mean Dice score) and TH intensity prediction accuracy (assessed via Smooth L1 loss). Task weights for segmentation and profiling losses were also adjusted to ensure balanced optimization during multi-task training. Learnable weights within the segmentation loss function (λ_1, λ_2) were initialized uniformly and jointly optimized with model parameters.The search space and selected ranges are summarized in Table 2.

Table 2. Hyperparameter Grid Used for Model Tuning

Hyperparameter	Values Explored	Selected Value
Initial learning rate	$1 \times 10^{-3}, 5 \times 10^{-4}, 1 \times 10^{-4}, 5 \times 10^{-5}$	1×10^{-4}
Task loss weights (α, β)	0.5-1.0	1.0
Dice vs. Focal loss weights (λ_1, λ_2)	Initialized as 0.5, learned during training	Learned dynamically
Dropout rate (refinement blocks)	0.0, 0.1, 0.2	0.1
Batch size	4, 8, 16	8
Scheduler type	Step decay, cosine annealing	Cosine annealing
Number of folds	5,10	5
Selection criterion	Mean Dice score, Smooth L1 loss	Combined metric

5.2. Performance Metrics

To comprehensively assess the performance of SegDAFormer, task-specific evaluation metrics were employed to measure accuracy across segmentation, neuron counting, and TH phenotypic profiling. These metrics were chosen to reflect both the structural fidelity of the predicted neuron boundaries and the biological relevance of intensity-based phenotype estimation.

The segmentation performance of SegDAFormer was assessed using four widely recognized metrics—Dice score, Precision, Recall, and F1 score—as defined in equations (33) to (36). These metrics

were computed based on the counts of True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN), derived from model predictions on the test dataset.

- Dice Score (F1 overlap): Measures the overlap between predicted and ground truth masks, defined as in (33). It reflects how well the predicted neuron somas align with the ground truth annotations, providing a direct indication of the model's accuracy in capturing the shape and extent of dopaminergic cell boundaries.

$$\text{Dice} = \frac{2TP}{2TP+FP+FN} \tag{33}$$

- Precision reflects the model's ability to avoid false detections, ensuring that the detected somas truly correspond to TH-positive neurons in the ground truth.

$$\text{Precision} = \frac{TP}{TP+FP} \tag{34}$$

- Recall measures the proportion of true somas that were correctly identified, highlighting the model's sensitivity and its ability to minimize missed detections.

$$\text{Recall} = \frac{TP}{TP+FN} \tag{35}$$

- The F1 score provides a balanced assessment of the model's performance by combining both precision and recall. It is especially useful in this context to evaluate how well SegDAFormer detects dopaminergic neurons while avoiding both false positives and missed detections, ensuring reliable segmentation across varying tissue conditions.

$$\text{F1} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision}+\text{Recall}} \tag{36}$$

To evaluate the model's effectiveness in estimating neuron counts, the following metrics were used:

Percentage Counting Error (PCE) measures the relative deviation from ground truth counts as in (37), where $\hat{N}$ is the predicted count and N is the true count. PCE shows how far the model's estimated number of dopaminergic neurons deviates from the actual count. It provides a simple way to understand how accurate the model is at matching the true number of neurons present in the tissue.

$$PCE = \left|\frac{\hat{N}-N}{N}\right| \times 100 \tag{37}$$

The Coefficient of Determination (R^2) assesses the correlation between predicted and actual counts as in (38). It indicates how well the model's predicted neuron counts track with the actual ground

truth across the dataset. A higher R^2 value suggests that the model captures the overall trend in neuron distribution accurately, reflecting strong agreement between predicted and true counts.

$$R^2 = 1 - \frac{\sum(N - \hat{N})^2}{\sum(N - \overline{N})^2} \tag{38}$$

To evaluate the model's ability to quantify dopaminergic phenotype, predicted TH intensity values were compared against ground truth measurements using regression-based metrics. These metrics assess how accurately the model captures the underlying biochemical state of neurons, beyond structural segmentation.

Mean Absolute Error (MAE) measures the average difference between the predicted and actual TH intensity values as in (39). It provides a straightforward indication of how close the model's estimates are to ground truth on a per-region or per-pixel basis. In this context, lower MAE values suggest that the model reliably captures the phenotypic expression levels of dopaminergic neurons.

$$\text{MAE} = \frac{1}{M}\sum_{j=1}^{M} |\hat{r_j} - r_j| \tag{39}$$

Root Mean Squared Error (RMSE) measures the square root of the average squared differences between predicted and actual TH intensity values as in (40), emphasizing larger errors. It complements MAE by being more sensitive to outliers, making it useful for detecting occasional large deviations in prediction. In this study, a lower RMSE indicates that the model's TH intensity estimates remain closely aligned with ground truth, even in regions with staining variability or complex neuronal structures. In these metrics, $\hat{r_j}$ and r_j are the predicted and ground truth TH intensity values, and are the predicted and ground truth TH intensity values, and M is the number of predicted regions or pixels.

$$\text{RMSE} = \sqrt{\frac{1}{M}\sum_{j=1}^{M} (\hat{r_j} - r_j)^2} \tag{40}$$

5.3. Performance Evaluation

This section presents the experimental results obtained for the proposed SegDAFormer model, evaluated on the test dataset using the metrics described in Section 5.2. The objective of this evaluation is to assess the model's ability to perform accurate segmentation of dopaminergic neuron somas and reliable prediction of their phenotypic state based on TH intensity. Results are reported as mean values with standard deviations across five cross-validation folds to reflect the model's consistency and generalization.

Table 3 summarizes the segmentation performance of SegDAFormer. The model achieved a mean Dice score of $93.12 \pm 1.08\%$, indicating a high degree of spatial overlap between predicted and ground truth neuron masks. Precision and recall were measured at $94.57 \pm 1.32\%$ and $91.98 \pm 1.26\%$, respectively, reflecting the model's ability to detect most true neuron regions while minimizing false

positives. The resulting F1 score of $93.24 \pm 1.17\%$ confirms that SegDAFormer maintains a strong balance between sensitivity and specificity.

These results highlight the effectiveness of the hierarchical encoder and MSIM in capturing detailed morphological features, while the edge-aware refinement and dual-head decoding architecture contribute to accurate boundary delineation.

Table 3. Segmentation Performance of SegDAFormer (Mean ± SD across 5 folds)

Metric	Value (%)
Dice Score	93.12 ± 1.08
Precision	94.57 ± 1.32
Recall	91.98 ± 1.26
F1 Score	93.24 ± 1.17

In addition to segmentation accuracy, the effectiveness of SegDAFormer in estimating the number of dopaminergic neurons was assessed using PCE and the R^2 , as shown in Table 4.

Table 4. Neuron Counting Performance of SegDAFormer

Metric	Value
PCE	5.82 ± 1.45
R^2	0.962 ± 0.014

The model achieved a mean counting error of $5.82 \pm 1.45\%$, demonstrating its ability to produce soma masks that align well with the biological ground truth in terms of count. Moreover, the high R^2value of 0.962 ± 0.014indicates a strong linear correlation between predicted and actual neuron counts, confirming that the model preserves quantitative consistency across the test dataset.

These findings suggest that SegDAFormer not only identifies neuron boundaries accurately but also enables reliable quantification of neuronal populations, a critical requirement for assessing neurodegeneration in preclinical Parkinson's disease models.The phenotypic profiling capability of SegDAFormer was evaluated by comparing predicted TH intensity values against expert-annotated ground truth using two regression metrics: MAE and RMSE. These metrics capture how closely the model's predictions align with biologically relevant intensity signals associated with dopaminergic neuron function.

As shown in Table 5, SegDAFormer achieved a mean MAE of 7.96 ± 1.22 and an RMSE of 11.43 ± 1.84, computed on an 8-bit grayscale intensity scale (0–255). These low error values demonstrate

the model's ability to capture subtle phenotypic gradients, even in regions with heterogeneous staining patterns or phenotypic downregulation.

This result confirms that the profiling head, supported by multi-scale features and masked pooling, enables accurate regression of biochemical phenotype while preserving spatial specificity—offering an interpretable and biologically meaningful output that complements structural segmentation.

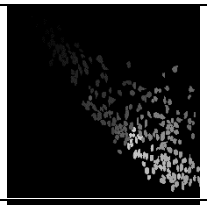
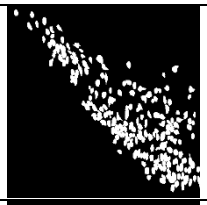
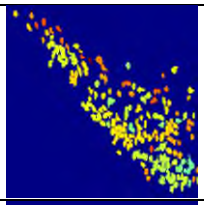
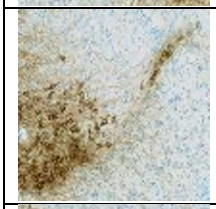
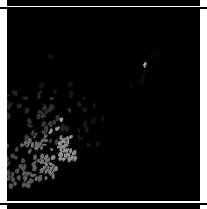
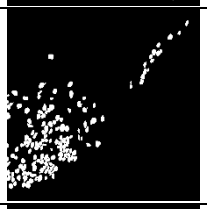
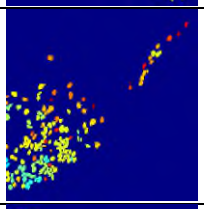
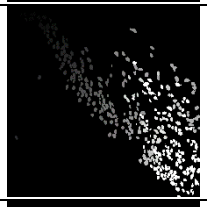
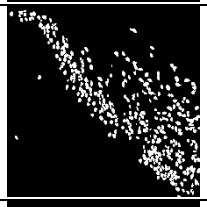
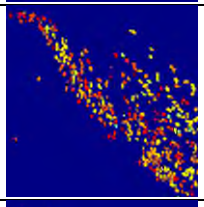
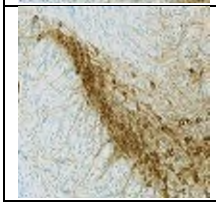
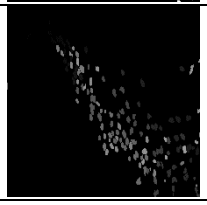
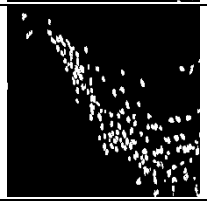
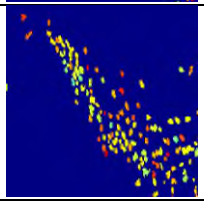
Table 5. TH Intensity Profiling Performance of SegDAFormer

Metric	Value
MAE %	7.96 ± 1.22
RMSE %	11.43 ± 1.84

Further, Table 6 presents representative outputs and quantitative performance metrics for SegDAFormer across four TH-stained histological images. For each sample, both segmentation quality and phenotypic profiling accuracy are assessed. The consistent and robust performance of SegDAFormer across varied histological sections is evident from this table. The Dice scores range from 91.4% to 93.5%, indicating high spatial overlap between predicted and ground truth neuron somas, even in anatomically complex or weakly stained regions. The best segmentation result is observed in Sample 1 ($93.5 \pm 1.0\%$), likely due to clear somatic boundaries and moderate cell density, while Sample 4 shows slightly lower performance ($91.4 \pm 1.3\%$), potentially reflecting challenges in boundary separation within densely clustered neurons.

Neuron counting accuracy, measured by the $\boldsymbol{R^2}$, follows a similar trend. Sample 1 achieves the highest correlation with manual counts (0.935 ± 0.02), affirming the model's ability to preserve quantitative integrity in well-defined regions. In contrast, Sample 4 presents a modest drop in $\boldsymbol{R^2}$ (0.902 ± 0.04), though still well above the threshold for reliable biological inference.Interestingly, phenotypic profiling accuracy, measured by MAEdoes not decline in proportion to segmentation quality. Sample 4, despite having the lowest Dice and $\boldsymbol{R^2}$, yields the lowest MAE ($4.3 \pm 0.7\%$), suggesting that the model can still extract meaningful TH intensity features even in morphologically ambiguous areas. This resilience is likely due to the profiling head's use of masked pooling and multi-scale context, which selectively aggregates phenotypic signals from neuron-positive regions.

Table 6. Sample-wise Evaluation of SegDAFormer on Histological Sections: Segmentation and Phenotypic Profiling Accuracy

Histology Image	Ground Truth	Predicted Segmentation Mask	Intensity Map	Dice	R^2	MAE%
				93.5 ± 1.0	0.935 ± 0.02	5.4 ± 1.0
				92.3 ± 1.2	0.922 ± 0.03	4.7 ± 0.8
				92.1 ± 0.9	0.914 ± 0.02	5.1 ± 0.9
				91.4 ± 1.3	0.902 ± 0.04	4.3 ± 0.7

Empirical evaluations confirm that SegDAFormer delivers robust and accurate performance across the key tasks of dopaminergic neuron segmentation, counting, and phenotypic profiling. The model consistently demonstrates high spatial precision, strong correlation with ground truth counts, and low error in TH intensity estimation. Its ability to generalize across diverse histological presentations highlights the effectiveness of its dual-head architecture, multi-scale feature integration, and phenotypically aware design. These results position SegDAFormer as a reliable and scalable framework for high-throughput analysis in preclinical studies of Parkinson's disease.

Qualitative inspection shows that pixel-wise TH expression is faithfully rendered across heterogeneous tissue regions. Warmer hues are consistently observed to coincide with areas of dense dopaminergic soma packing—particularly within the dorsal tier of the substantia nigra—while cooler tones appear in zones of reduced staining intensity, such as lesion-affected regions. Smooth color gradations at the interfaces between high- and low-expression domains demonstrate that subtle phenotypic transitions are delineated without introducing spurious artifacts. Even in regions of closely apposed or faintly stained cells, clear contrast is maintained, indicating that both local detail and broader

context are preserved through multi-scale feature fusion. Overall, these visual representations align with known anatomical distributions and lesion-induced TH loss, thereby validating the model's capacity to generate biologically interpretable intensity maps.

5.4. Comparative Analysis

SegDAFormer's performance was evaluated against several leading models in dopaminergic neuron analysis, all tested on the same TH-stained dataset. Table 7 presents a comparative summary of SegDAFormer and existing models across key performance metrics. The representative models for comparison include multiclass anatomical segmentation model by Barzekar et al. [22], general-purpose instance segmentation using Cellpose by Stringer et al. [12], customized CNN-based detection by Kim et al. [23], categorical intensity estimation via TrueTH by Chen et al. [25], post hoc profiling using Swin U-Net by Wang et al. [31], and Transformer-based subregion segmentation from Nebel et al. [29]. The comparison spans segmentation accuracy, neuron count estimation, and TH phenotype profiling performances.

SegDAFormer outperforms these models, with a Dice score of 93.12% and the lowest counting error (5.82%) among all methods. This improvement stems from its Hierarchical Transformer Encoder, which captures long-range spatial dependencies overlooked by traditional CNNs [12, 23]. The MSIM plays a critical role in harmonizing encoder features, contributing to superior boundary precision over methods like Barzekar et al. [22] and Nebel et al. [29], which focus primarily on anatomical region segmentation. Unlike TrueTH [25] and Swin U-Net [31], which either provide coarse intensity categories or rely on post hoc inference, SegDAFormer introduces a dedicated phenotypic profiling head that directly regresses TH intensity using masked, structure-aware pooling and uncertainty-guided refinement, yielding 0.962 .

Table 7. Comparative Evaluation of SegDAFormer with State-of-the-art

Model	Dice (%)	Counting Error (%)	R^2	Phenotype Profiling	Limitations / Inference
Multiclass Segmentation [22] Barzekar et al. (2023)	87.42	10.56	0.905	Partial (region-based)	Focuses on anatomical regions; lacks per-cell profiling granularity.
Cellpose [12] Stringer et al. (2021)	85.72	9.91	0.893	No	General-purpose; not optimized for TH neuron characteristics.
Customized CNN [23] Kim et al. (2023)	84.22	12.35	0.876	No	Designed for detection and counting only; lacks segmentation and profiling capability.
TrueTH [25] Chen et al. (2024)	88.43	9.08	0.917	Partial (categorical)	Provides only intensity categories; lacks fine-grained detail.

Model	Dice (%)	Counting Error (%)	R^2	Phenotype Profiling	Limitations / Inference
Swin U-Net [31] Wang et al. (2024)	85.60	10.67	0.884	Weak (post hoc only)	No integrated profiling; requires external processing.
Conv-SeTr [29] Nebel et al. (2023)	87.15	10.14	0.889	—	—
SegDAFormer (Proposed)	**93.12**	**5.82**	**0.962**	**Yes (direct + multiscale)**	Integrates structure and phenotype; accurate and interpretable.

SegDAFormer outperforms existing models in segmentation, counting, and phenotypic profiling, owing to its integrated architecture and biologically informed design. Its dual-head framework enables accurate, interpretable analysis, setting a new benchmark for automated dopaminergic neuron assessment.

5.5. Ablation Study

To understand the contribution of individual architectural components to the overall performance of SegDAFormer, an ablation study was conducted using systematically modified versions of the model. Each variant was designed by selectively removing or altering a key module while keeping all other components unchanged. The evaluation was carried out using the same test set and metrics as the full model, allowing for direct comparison. This analysis offers insight into the functional significance of multi-scale integration, edge refinement, phenotype modeling, and structural context in the joint segmentation and profiling tasks.

The following variants were tested: (A1) the full SegDAFormer model, (A2) removal of the MSIM, (A3) exclusion of edge attention in the segmentation head, (A4) removal of the phenotypic profiling head, (A5) removal of the confidence estimation branch, and (A6) replacement of masked pooling with global average pooling in the profiling head.

The ablation study presented in Table 8 quantitatively evaluates the impact of removing individual components from the SegDAFormer architecture. Each variant was assessed on the full test set using segmentation and profiling metrics. The full model (A1) consistently achieved the best results across all tasks, confirming the efficacy of its integrated design. In contrast, the ablated variants (A2–A6) exhibited varying degrees of performance degradation, ranging from modest to substantial, depending on the removed module.

The removal of the MSIM (A2) resulted in a notable drop in segmentation performance (Dice: 83.81 ± 1.56) and a 10% increase in MAE and RMSE, highlighting the importance of hierarchical multi-resolution fusion for both boundary localization and phenotypic accuracy. Without MSIM, the encoder features were insufficiently harmonized, leading to noisy predictions and reduced spatial precision.

A3, which excludes the edge refinement layer, showed a further drop in Dice (81.95 ± 1.48) and an increase in TH estimation error (MAE: 8.91 ± 1.41), indicating that boundary-aware decoding plays a

critical role in accurately separating closely packed neurons and in improving phenotype alignment at the edges.

A4 disabled the phenotypic profiling head entirely. While segmentation metrics (Dice: 79.15 ± 1.72) were preserved at a basic level, the model failed to generate any phenotypic outputs, confirming that structural segmentation alone is insufficient for comprehensive histological interpretation.

The absence of confidence estimation (A5) led to a broader decline in both segmentation and profiling metrics. The RMSE increased by more than 20%, and $\boldsymbol{R^2}$ dropped to 0.770 ± 0.026, indicating that the uncertainty branch contributes to stabilizing predictions in low-contrast or ambiguous regions by attenuating noisy regressions. Lastly, the removal of masked pooling (A6) resulted in the most significant degradation across all metrics (Dice: 69.84 ± 2.02; MAE: 9.95 ± 1.52; RMSE: 14.29 ± 2.39), confirming that the use of structure-guided pooling is essential for accurate TH intensity estimation. Without it, the profiling head incorporated features from background or non-neuronal areas, reducing biological specificity and increasing error variance.

This ablation study validates the functional necessity of each architectural component in SegDAFormer. It reveals that the MSIM and masked pooling modules are the most critical for preserving segmentation accuracy and phenotypic precision, while edge refinement and confidence estimation further enhance robustness and interpretability. These findings demonstrate the integrity of the model's modular design and its ability to deliver reliable dual-task performance in histological analysis.

Table 8. Ablation Study Results for SegDAFormer Variants

Variant	Dice (%)	F1 Score (%)	Counting Error (%)	R^2	MAE (TH)	RMSE (TH)
A1: Full Model	93.12 ± 1.08	93.24 ± 1.17	5.82 ± 1.45	0.962 ± 0.014	7.96 ± 1.22	11.43 ± 1.84
A2: No MSIM	83.81 ± 1.56	83.92 ± 1.63	6.40 ± 1.62	0.866 ± 0.019	8.76 ± 1.38	12.86 ± 2.01
A3: No Edge Ref.	81.95 ± 1.48	82.05 ± 1.55	6.52 ± 1.58	0.847 ± 0.021	8.91 ± 1.41	12.80 ± 1.97
A4: No Profiling Head	79.15 ± 1.72	79.25 ± 1.68	6.69 ± 1.71	0.818 ± 0.023	—	—
A5: No Conf. Est.	74.50 ± 1.89	74.59 ± 1.91	6.98 ± 1.77	0.770 ± 0.026	9.55 ± 1.45	13.72 ± 2.24
A6: No Masked Pooling	69.84 ± 2.02	69.93 ± 2.05	7.28 ± 1.86	0.722 ± 0.029	9.95 ± 1.52	14.29 ± 2.39

5.6. Explainable Analysis

Interpretability plays a vital role in bridging the gap between computational predictions and biological validation, especially in histopathological applications. Beyond quantitative performance, it is essential to understand how and why the model arrives at its predictions. SegDAFormer addresses this

need by producing spatially resolved outputs that can be directly compared to known biological patterns, allowing researchers to assess model reliability in real tissue contexts.

Figure 6 provides a clear example of how SegDAFormer's predictions can be meaningfully interpreted alongside the original histological images. The top row shows TH-stained midbrain tissue, while the bottom row displays the model's TH intensity predictions overlaid on the segmented dopaminergic neurons. The colors in the overlay reflect the predicted level of TH expression, while brighter colors indicate healthier neurons with stronger TH signals, cooler tones suggest lower expression, which may reflect early signs of stress or degeneration. This allows researchers to not only see where the neurons are but also assess their phenotypic state at a glance.

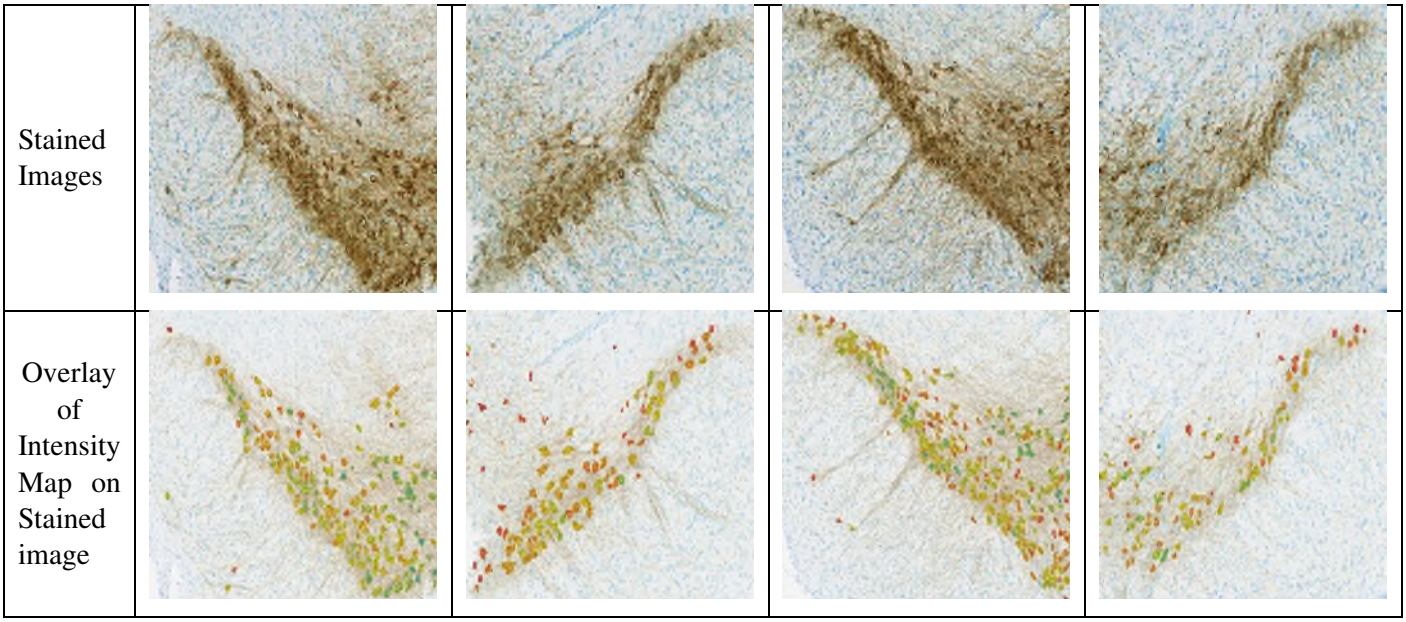


Figure 6. Explainable analysis of SegDAFormer

By combining structural segmentation with continuous biochemical profiling, SegDAFormer enables intuitive interpretation of subtle variations in dopaminergic neuron health. Its outputs align closely with biological patterns, offering a transparent and practical tool for studying tissue changes in Parkinson's disease.

5.7. Discussion

The development of SegDAFormer is a significant advancement in the automated analysis of dopaminergic neurons, particularly for histological studies in Parkinson's disease. By integrating segmentation and phenotypic profiling into a unified, Transformer-based framework, the model directly

addresses long-standing limitations in existing methods that either focus exclusively on structure or rely on disconnected pipelines for biochemical inference.

Quantitatively, SegDAFormer demonstrated strong generalization across high-resolution TH-stained mouse brain images, achieving a Dice score of $93.12 \pm 1.08\%$, a low counting error of $5.82 \pm 1.45\%$, and a high correlation with ground truth counts($\boldsymbol{R^2}$= 0.962R2). These results exceed the performance of state-of-the-art alternatives, including U-Net-based architectures and recent transformer-CNN hybrids. The model's phenotypic profiling also proved robust, with a MAE of 7.96 ± 1.22 and an RMSE of 11.43 ± 1.84, supporting accurate regression of TH intensity at both pixel and regional levels. These metrics validate SegDAFormer's capacity to capture both morphometric precision and biochemical relevance.

Further, explainable overlays illustrated how model predictions align with biological staining patterns, offering direct insight into phenotypic heterogeneity within dopaminergic populations. By visualizing TH expression gradients through color-coded maps superimposed on neuron somas, the model enables researchers to distinguish intact neurons from those exhibiting biochemical downregulation or early phenotypic deterioration. This interpretability is especially valuable in preclinical research, where assessing not only neuron counts but their viability is essential for evaluating disease progression or therapeutic intervention.

The model's architecture, built around a hierarchical Transformer encoder, the MSIM, and dual decoding headswas key to its overall performance. However, noticeable drops in performance were observed in the ablation study when components such as MSIM, edge refinement, or the phenotypic profiling head were removed. The profiling head was identified as especially critical. By integrating pixel-level intensity estimation with region-aware, multi-scale feature fusion, subtle variations in TH expression across the tissue were captured. This combination ensured that outputs were both accurate and easily interpreted.

SegDAFormer was shown to have broad utility in neuroscience and preclinical research: in Parkinson's disease models, dopaminergic neuron loss was automatically quantified and temporal changes in TH staining were detected, providing a more rapid and scalable alternative to manual counting. Because both neuronal structure and biochemical state were jointly profiled, the effects of therapies aimed at preserving or restoring dopaminergic function could be more precisely assessed. Adaptation of the framework for other neurodegenerative disorders—such as Alzheimer's or Huntington's disease—was deemed feasible, and with further training, support for multiplex histology (e.g. co-expression analysis of VMAT2 or DAT) could be enabled, yielding a more comprehensive view of neuronal identity and function.

Further, SegDAFormer could be integrated into digital pathology platforms to support high-throughput, reproducible analysis of brain tissue in both experimental and clinical settings.

Although strong performance and biological relevance have been demonstrated by SegDAFormer, certain limitations should be acknowledged. Accuracy may be degraded in regions with poor staining quality, high background noise, or atypical neuronal morphologies (e.g., fragmentation or extreme crowding), compromising both segmentation and phenotypic profiling outputs and resulting in neurons being undercounted or misclassified. Furthermore, reliance on TH as the sole phenotypic marker may fail to capture the full heterogeneity of dopaminergic neuron function. To enable more comprehensive cellular profiling, additional markerssuch as GFAP, NeuN, or oxidative stress indicatorscould be incorporated in future work. Integration of temporal datasets is proposed to support longitudinal analysis of neuronal changes. Finally, generalization across diverse staining protocols, species, and tissue preparations could be improved through the adoption of weakly supervised or semi-supervised learning strategies, thereby broadening the model's applicability in both experimental and clinical contexts.

SegDAFormer offers a robust and interpretable solution for automated analysis of dopaminergic neurons, bridging the gap between structural segmentation and biochemical insight. Through its dual-head architecture and biologically informed design, it enables researchers to move beyond counting cells toward understanding their functional state within complex tissue environments. By combining technical precision with scientific relevance, SegDAFormer lays the groundwork for scalable, reproducible, and meaningful histological analysis—opening new avenues for discovery in Parkinson's disease and related neurodegenerative research.

6. CONCLUSION

This study addresses a fundamental limitation in the automated analysis of dopaminergic neurons—the disconnect between structural segmentation and biochemical profiling. In this research, this gap is bridged by the introduction of SegDAFormer, an end-to-end Transformerbased framework that unites soma-level segmentation and pixel-wise TH intensity profiling within a single, interpretable model. A dual-headed design, supported by a hierarchical feature fusion module, has been employed to blend fine morphological detail with biochemical signals, thereby capturing subtle phenotypic gradients of neuronal health. Through comprehensive benchmarking against established baselines, both segmentation accuracy and TH quantification precision have been shown to surpass current standards, and transparency in decision making has been demonstrated by integrated explainability mechanisms. This advance is expected to transform preclinical workflows in Parkinson's disease research by enabling rapid, large-scale phenotypic analysis.

The modular architecture of this model facilitates straightforward extension to multi-marker histologyincluding GFAP, NeuN, and oxidative stress indicators. In future research, the annotation burden will be alleviated through integration of semi-supervised and active-learning strategies. Robustness to staining and preparation variability will be enhanced by implementing normalization and augmentation pipelines that are agnostic to such differences. Temporal coherence across sequential sections will be strengthened by embedding continuity constraints within the model's training objectives. Ultimately, it is envisioned that SegDAFormer will be established as a cornerstone of digital pathology, uniting computational precision with biological insight to accelerate discoveries in neuroscience.

Author contributions

H. M. S. & R. S.: Reviewed overall work, planning, and project design.

H. M. S. & G. P. A. M.: Contributed to methodology implementation and core development.

G. P. A. M. & M. N. G.: Assisted in data preparation, system development, and documentation.

M. N. G. & M.T.R.: Participated in formal analysis, validation, and content organization.

Data Availability Statement

The dataset for the SegDAFormer paper is available at https://data.mendeley.com/datasets/8phmy565nk/1 for download [38] in the paper.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declarations

Conflict of Interest

The authors declare that they have no competing interests related to this work.

Ethics Approval and Consent to Participate

Not applicable. The dataset used is publicly available and does not involve any personally identifiable information or direct participation of human subjects.

Consent for Publication

Not applicable, as no personal or sensitive data was used in the study.

Clinical Trial Registration

References

1. Poewe, W., Seppi, K., Tanner, C. M., Halliday, G. M., Brundin, P., Volkmann, J., ... & Lang, A. E. (2017). Parkinson disease. *Nature reviews Disease primers*, *3*(1), 1-21.
2. Dauer, W., & Przedborski, S. (2003). Parkinson's disease: mechanisms and models. *Neuron*, *39*(6), 889-909.
3. Guatteo, E., Berretta, N., Monda, V., Ledonne, A., & Mercuri, N. B. (2022). Pathophysiological features of nigral dopaminergic neurons in animal models of Parkinson's disease. *International journal of molecular sciences*, *23*(9), 4508.
4. Daubner, S. C., Le, T., & Wang, S. (2011). Tyrosine hydroxylase and regulation of dopamine synthesis. *Archives of biochemistry and biophysics*, *508*(1), 1-12.
5. Kawahata, I., & Fukunaga, K. (2020). Degradation of tyrosine hydroxylase by the ubiquitin-proteasome system in the pathogenesis of Parkinson's disease and dopa-responsive dystonia. *International journal of molecular sciences*, *21*(11), 3779.
6. Pajarillo, E., Rizor, A., Son, D. S., Aschner, M., & Lee, E. (2020). The transcription factor REST up-regulates tyrosine hydroxylase and antiapoptotic genes and protects dopaminergic neurons against manganese toxicity. *Journal of Biological Chemistry*, *295*(10), 3040-3054.
7. Dauer, W., & Przedborski, S. (2003). Parkinson's disease: mechanisms and models. *Neuron*, *39*(6), 889-909.
8. Baquet, Z. C., Williams, D., Brody, J., & Smeyne, R. (2009). A comparison of model-based (2D) and design-based (3D) stereological methods for estimating cell number in the substantia nigra pars compacta (SNpc) of the C57BL/6J mouse. *Neuroscience*, *161*(4), 1082-1090.
9. West, M. J., Slomianka, L. H. J. G., & Gundersen, H. J. G. (1991). Unbiased stereological estimation of the total number of neurons in the subdivisions of the rat hippocampus using the optical fractionator. *The Anatomical Record*, *231*(4), 482-497.
10. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18* (pp. 234-241). Springer international publishing.
11. He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision* (pp. 2961-2969).
12. Stringer, C., Wang, T., Michaelos, M., & Pachitariu, M. (2021). Cellpose: a generalist algorithm for cellular segmentation. *Nature methods*, *18*(1), 100-106.
13. Caron, M., Misra, I., Mairal, J., Goyal, P., Bojanowski, P., & Joulin, A. (2020). Unsupervised learning of visual features by contrasting cluster assignments. *Advances in neural information processing systems*, *33*, 9912-9924.
14. Zbontar, J., Jing, L., Misra, I., LeCun, Y., & Deny, S. (2021, July). Barlow twins: Self-supervised learning via redundancy reduction. In *International conference on machine learning* (pp. 12310-12320). PMLR.
15. Ip, C. W., Cheong, D., & Volkmann, J. (2017). Stereological estimation of dopaminergic neuron number in the mouse substantia nigra using the optical fractionator and standard microscopy equipment. *Journal of visualized experiments: JoVE*, (127), 56103.
16. Fabricius, K., Barkholt, P., Jelsing, J., & Hansen, H. H. (2017). Application of the physical disector principle for quantification of dopaminergic neuronal loss in a rat 6-hydroxydopamine Nigral lesion model of Parkinson's disease. *Frontiers in neuroanatomy*, *11*, 109.

17. Tian, Y., Johnson, G. A., Williams, R. W., & White, L. E. (2023). A rapid workflow for neuron counting in combined light sheet microscopy and magnetic resonance histology. *Frontiers in Neuroscience*, *17*, 1223226.
18. Kim, D., Bak, M. S., Park, H., Baek, I. S., Chung, G., Park, J. H., ... & Kim, S. K. (2023). An automated cell detection method for th-positive dopaminergic neurons in a mouse model of Parkinson's disease using convolutional neural networks. *Experimental Neurobiology*, *32*(3), 181.
19. Penttinen, A. M., Parkkinen, I., Blom, S., Kopra, J., Andressoo, J. O., Pitkänen, K., ... & Airavaara, M. (2018). Implementation of deep neural networks to count dopamine neurons in substantia nigra. *European Journal of Neuroscience*, *48*(6), 2354-2361.
20. Brewer, B. R., Pradhan, S., Carvell, G., & Delitto, A. (2009). Application of modified regression techniques to a quantitative assessment for the motor signs of Parkinson's disease. *IEEE transactions on neural systems and rehabilitation engineering*, *17*(6), 568-575.
21. Falk, T., Mai, D., Bensch, R., Çiçek, Ö., Abdulkadir, A., Marrakchi, Y., ... & Ronneberger, O. (2019). U-Net: deep learning for cell counting, detection, and morphometry. *Nature methods*, *16*(1), 67-70.
22. Barzekar, H., Ngu, H., Lin, H. H., Hejrati, M., Valdespino, S. R., Chu, S., ... & Ghosh, S. (2023). Multiclass Semantic Segmentation to Identify Anatomical Sub-Regions of Brain and Measure Neuronal Health in Parkinson's Disease. *arXiv preprint arXiv:2301.02925*.
23. Kim, D., Bak, M. S., Park, H., Baek, I. S., Chung, G., Park, J. H., ... & Kim, S. K. (2023). An automated cell detection method for th-positive dopaminergic neurons in a mouse model of Parkinson's disease using convolutional neural networks. *Experimental Neurobiology*, *32*(3), 181.
24. Pandey, P. M., Mitra, S., Singh, S., & Singh, A. (2024). Deciphering Parkinson's Disease Progression Using Deep Learning Techniques. *SN Computer Science*, *5*(8), 1042.
25. Chen, J., Meng, Q., Zhang, Y., Liang, Y., Ding, J., Xia, X., & Hu, G. (2024). TrueTH: A user-friendly deep learning approach for robust dopaminergic neuron detection. *Neuroscience Letters*, *836*, 137871.
26. Wang, M., Xiao, L., Shi, Y., Wu, Y., Huang, X., Wu, Y., ... & Wang, W. (2025). Whole-brain 3D imaging of dopaminergic neurons and glial cells in the mouse model of Parkinson's disease induced by 6-OHDA. Frontiers in Aging Neuroscience, 17, 1503168.
27. Fred, A. L., Kumar, A., Elias, T. S., & RF, F. R. (2024, December). Substantia Nigra Segmentation in 7T MRI Using a 3D UNet Architecture for Enhanced Parkinson's Disease Diagnosis. In *2024 International Conference on Brain Computer Interface & Healthcare Technologies (iCon-BCIHT)* (pp. 123-127). IEEE.
28. Basukala, D., Mukundan, R., Lim, A., Hurrell, M. A., Keenan, R. J., Dalrymple-Alford, J. C., ... & Melzer, T. R. (2021). Automated segmentation of substantia nigra and red nucleus using quantitative susceptibility mapping images: Application to Parkinson's disease. *Computers & electrical engineering*, *91*, 107091.
29. Nebel, J., Pinargote, F. E. M., Peláez, C. E., Paredes, F. R. L., & Rodriguez-Rojas, R. (2023, April). Subthalamic Nucleus and Substantia Nigra Automatic Segmentation Using Convolutional Segmentation Transformers (Conv-SeTr). In *World Conference on Information Systems and Technologies* (pp. 357-367). Cham: Springer Nature Switzerland.
30. Haghighi, F., Ghosh, S., Ngu, H., Chu, S., Lin, H., Hejrati, M., ... & Hashemifar, S. (2023). Self-supervised Learning for Segmentation and Quantification of Dopamine Neurons in Parkinson's Disease. *arXiv preprint arXiv:2301.08141*.
31. Wang, T., Wang, Y., Zhu, H., Liu, Z., Chen, Y. C., Wang, L., ... & Jiang, L. (2024). Automatic substantia nigra segmentation with Swin-Unet in susceptibility-and T2-weighted imaging: application to Parkinson disease diagnosis. *Quantitative Imaging in Medicine and Surgery*, *14*(9), 6337.

32. Hussain, S. S., Shah, P. M., Dawood, H., Degang, X., Alshamayleh, A., Khan, M. A., & Ghazal, T. M. (2025). A swin transformer and CNN fusion framework for accurate Parkinson disease classification in MRI. *Scientific Reports*, *15*(1), 15117.

33. Lu, M., Christensen, C. N., Weber, J. M., Konno, T., Läubli, N. F., Scherer, K. M., ... & Kaminski, C. F. (2023). ERnet: a tool for the semantic segmentation and quantitative analysis of endoplasmic reticulum topology. *Nature Methods*, *20*(4), 569-579.

34. Di Credico, A., Weiss, A., Corsini, M., Gaggi, G., Ghinassi, B., Wilbertz, J. H., & Di Baldassarre, A. (2023). Machine learning identifies phenotypic profile alterations of human dopaminergic neurons exposed to bisphenols and perfluoroalkyls. *Scientific Reports*, *13*(1), 21907.

35. D'Sa, K., Evans, J. R., Virdi, G. S., Vecchi, G., Adam, A., Bertolli, O., ... & Gandhi, S. (2023). Prediction of mechanistic subtypes of Parkinson's using patient-derived stem cell models. *Nature Machine Intelligence*, *5*(8), 933-946.

36. Weiss, A., Sommer, P., & Wilbertz, J. H. (2023). Phenotypic Profiling of Human Stem Cell-Derived Midbrain Dopaminergic Neurons. *Journal of Visualized Experiments (JoVE)*, (197), e65570.

37. Villaronga-Luque, A., Savill, R. G., López-Anguita, N., Bolondi, A., Garai, S., Gassaloglu, S. I., ... & Veenvliet, J. V. (2025). Integrated molecular-phenotypic profiling reveals metabolic control of morphological variation in a stem-cell-based embryo model. *Cell Stem Cell*, *32*(5), 759-777,

38. https://data.mendeley.com/datasets/8phmy565nk/1. [Accessed on 20 May 2025]