

HEDS-Net: Minimax-Based Segmentation-Guided Melanoma Detection from Dermoscopic Images

A. Kousar Nikhath¹, Sireesha Chittepu², M Swapna³, Rashmi Bhale⁴, Yeshwanth Reddy Peddamallu⁵

¹Assoc.Prof, Department of CSE (AIML &IoT),R&AI, Vallurupalli Nageswara Rao Vignana Jyothi Institute of Engineering and Technology, Vignana Jyothi Nagar, Bachupally, Hyderabad, Telangana, India. kousarnikhath@vnrvjiet.in

²Asst. Prof, Department of Information Technology, Vasavi College of Engineering, Ibrahimbagh, Hyderabad, Telangana, India. sireesha@staff.vce.ac.in

³Asst. Prof, Department of Computer Science and Engineering, Anurag University, Hyderabad, Telangana, India. mswapna.cse@anurag.edu.in

⁴Asst Prof, School of Engineering, Anurag University, Hyderabad, Telangana, India. rashmi.cse@anurag.edu.in

⁵AI Engineer, Infosys Pvt Ltd, Dallas, USA. peddamalluyeshwanthreddy@gmail.com

Abstract

Hair, ruler marks and marker ink corrupt lesion borders in dermoscopic images and mislead both segmentation and classification, while malignant lesions are scarce in public archives. This paper presents HEDS-Net (Hilbert-Edge, Difference-convolution and Structure-aware Network), a three-stage pipeline for melanoma detection. First, a Hilbert–Sobel artefact removal (HSAR) module detects thin artefacts from the difference between the directional analytic-signal amplitudes of the image and of an oriented-closing reference, inpaints them, restores colour and contrast, and outputs a Hilbert edge prior. Second, a difference-convolution re-parameterised segmentation decoder (DRS-Decoder) uses tri-branch difference convolutions and an edge-guided dispersion gate to delineate low-contrast borders, and folds every multi-branch kernel into a single 3×3 convolution at inference. Third, a knowledge-guided classifier fuses ABCD shape, colour and texture descriptors computed from the predicted mask with deep features of a two-phase fine-tuned ResNet-50. On ISIC-2018, HEDS-Net reaches a lesion Dice of 91.84%. On ISIC-2020 it attains 99.21% accuracy under the protocol of the SegFusion base framework (99.01%) and 92.31% under a leakage-free lesion-level protocol, where the re-run base framework obtains 89.74%.

Keyword: Dermoscopy, hair removal, Hilbert transform, difference convolution, skin lesion segmentation, melanoma classification, ResNet-50.

I. Introduction

Melanoma accounts for most skin-cancer deaths although it is a minority of skin-cancer cases [1], and its survival depends strongly on early detection. Dermatologists rely on structured rules such as the ABCD rule of dermoscopy, which weighs asymmetry, border, colour and differential structures [2]; automating such rules requires a clean image, an accurate lesion mask and a classifier that remains reliable when malignant cases are rare [3].

Current systems fall short on all three counts. Hair, ruler marks and ink overlap lesion borders, and classical morphological removal either leaves residual hair or erases genuine pigment [4]. Early lesions have faint borders that standard convolutions smooth away. Melanoma forms about 1.8% of ISIC-2020 [5], and image-level splits of augmented data can leak the same lesion into training and test sets and inflate accuracy [6].

The SegFusion framework [7] links a U-Net [8] trained on HAM10000 masks [9] with an EfficientNet-B0 [10] classifier through a mask-blending bridge and reports 99.01% accuracy on ISIC-2020. However, it applies no artefact removal, reports segmentation only as 91.2% pixel accuracy, uses no clinical descriptors, and splits the data after resampling and augmentation. The contributions of this paper are:

1. A Hilbert–Sobel artefact removal (HSAR) module that detects hair, ruler marks and ink and passes a Hilbert edge prior to the segmenter.
2. A difference-convolution re-parameterised segmentation decoder (DRS-Decoder) that sharpens low-contrast borders at the inference cost of plain convolution.
3. A knowledge-guided ResNet-50 [11] classifier that fuses mask-derived ABCD descriptors with deep features, evaluated against SegFusion under both its own protocol and a leakage-free lesion-level protocol.

II. Related Work

DullRazor [4] removes hair by grey-level closing and interpolation, and a deep encoder–decoder trained on synthetic hair improved removal quality [12], while a public dataset now provides real hair masks [13]. For segmentation, U-Net [8], TransUNet [14] and the lightweight EGE-UNet [15], which reaches a Dice of 89.46% on ISIC-2018 [16], are common baselines, and DualCoAg-Net combined contrastive aggregation with a Bayesian decoder in a federated multi-task network [17]. For classification, a diversity-driven CNN ensemble reached 90.15% on ISIC-2018 [18], a hybrid quantum–classical model fused deep and handcrafted descriptors [19], and SkinViT applied a transformer to melanoma classification [20]. SegFusion [7] showed that blending a predicted mask into the image improves EfficientNet-B0 classification. In these works the artefact map is not reused for segmentation, and mask-derived clinical descriptors are not combined with deep features under a leakage-free evaluation.

III. Proposed Methodology

Fig. 1 shows the three stages of HEDS-Net.

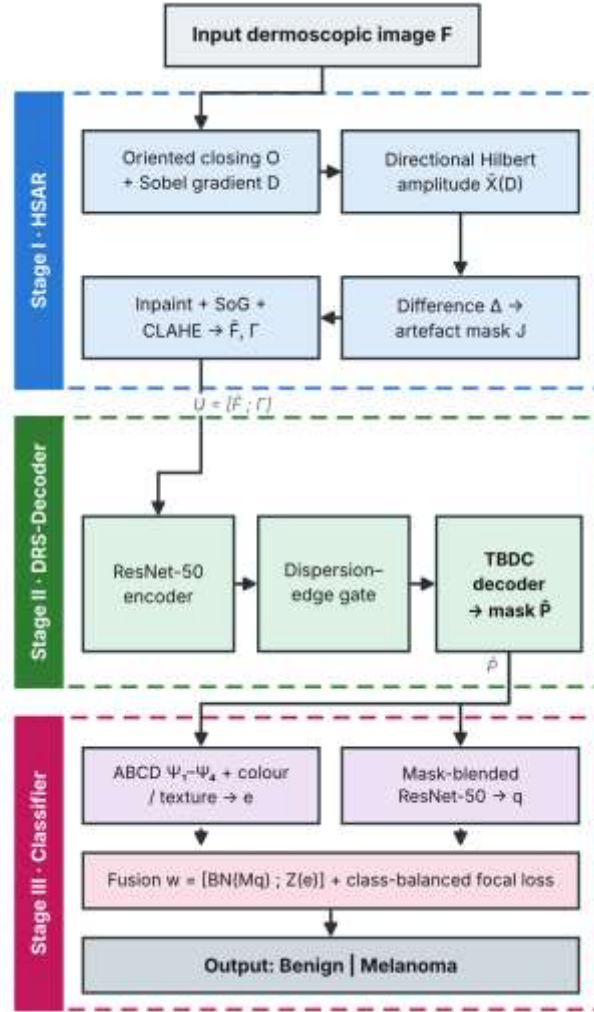


Fig. 1. Compact HEDS-Net framework: (I) Hilbert–Sobel artefact removal, (II) DRS-Decoder lesion segmentation, (III) knowledge-guided ResNet-50 classification.

A. Stage I: Hilbert–Sobel Artefact Removal (HSAR)

From the luminance I , a hair-suppressed reference O is obtained as the minimum of H oriented closings with linear structuring elements P_{θ_h} , and the Sobel gradient magnitude D_Q is computed for $Q \in \{I, O\}$:

$$O = \min_h (I \bullet P_{\theta_h}) \quad (1)$$

$$D_Q = \sqrt{(e_u * Q)^2 + (e_v * Q)^2} \quad (2)$$

where e_u and $e_v = e_u^\top$ are the Sobel kernels. The Hilbert transform $E\{\cdot\}$, defined with the Cauchy principal value P , is applied along each orientation θ_h to form a directional analytic signal whose mean amplitude measures oscillation energy:

$$E\{c\}(u) = \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{c(p)}{u-p} dp \quad (3)$$

$$X^-(D_Q) = \frac{1}{H} \sum_{h=1}^H |D_Q + i E_{\theta_h}\{D_Q\}| \quad (4)$$

Because O keeps the lesion but not thin structures, their amplitude difference isolates artefacts; Otsu thresholding with an elongation filter $\mathcal{E}$ gives the hair mask J_u , and a saturation–hue rule gives the ink mask J_v :

$$\Delta = |X^-(D_I) - X^-(D_O)| \quad (5)$$

$$J = \mathcal{E}(1[\Delta > \tau]) \vee J_v \quad (6)$$

The masked pixels are inpainted, Shades-of-Grey colour constancy and CLAHE are applied to obtain $\tilde{F}$, and the normalised Hilbert amplitude Γ of the restored image is concatenated as an edge prior, $U = \text{Concat}(\tilde{F}, \Gamma)$.

B. Stage II: DRS-Decoder

A ResNet-50 encoder with four input channels feeds a top-down decoder. Each decoder convolution is a tri-branch difference convolution (TBDC) that combines a vanilla kernel T_0 with central-difference and angular-difference kernels t_1 and t_2 over the 3×3 neighbourhood Ω_3 :

$$v_1(m_0) = \sum_{m_k \in \Omega_3} t_1(m_k)[u(m_0 + m_k) - u(m_0)] \quad (7)$$

$$v_2(m_0) = \sum_{m_k \in \Omega_3} t_2(m_k)[u(m_0 + m_k) - u(m_0 + m_k')] \quad (8)$$

where m_k' is the clockwise neighbour of m_k . The branches are fused with learnable weights during training and folded into one kernel at inference:

$$T_1'(m_k) = t_1(m_k) - \theta 1(m_k = 0) \sum_{g \in \Omega_3} t_1(g) \quad (9)$$

$$T^* = \alpha_0 T_0 + \alpha_1 T_1' + \alpha_2 T_2' \quad (10)$$

Before fusion, each skip feature C is refined by a dispersion–edge gate that weights channels by their mean μ and standard deviation σ and gates pixels with the feature–difference residual and the down-sampled edge prior $\Gamma_\downarrow$:

$$T_{ze} = \text{Sig}(\text{MLP}([\mu; \sigma])) \quad (11)$$

$$J_{pm} = \text{HardSig}(\text{Conv}_{1 \times 1}([C - \text{TBDC}(C); \Gamma_\downarrow])) \quad (12)$$

$$\text{DEG}(C) = (1 - \gamma) C + \gamma(C \otimes T_{ze} \otimes J_{pm}) \quad (13)$$

Four side outputs P_f are supervised against the mask G with $I_{seg} = \sum_f [I_{BCE} + I_{IoU} + \lambda_1 I_{Dice}](P_f, G)$.

C. Stage III: Knowledge-Guided Classifier

From the predicted mask $\hat{P}$, four ABCD descriptors are computed: asymmetry about the principal axes, border compactness from the perimeter Π , the number of the six dermoscopic colours present, and the maximum Feret diameter scaled by the pixel pitch q :

$$\psi_1 = \frac{1}{2} \sum_{\theta} \frac{|\hat{P} \oplus R_{\theta}(\hat{P})|}{|\hat{P}|}, \quad \psi_2 = \frac{\Pi^2}{4\pi|\hat{P}|} \quad (14)$$

$$\psi_4 = q \max_{u, w \in \partial \hat{P}} \|u - w\| \quad (15)$$

Together with the colour count ψ_3 , CIELab colour statistics, grey-level co-occurrence and local binary pattern texture, they form the knowledge vector e . The enhanced image is blended with its mask as in [7] and passed to a ResNet-50 whose pooled feature q is fused with e , and the network is trained with a class-balanced focal loss:

$$w = [\text{BN}(Mq); Z(e)] \quad (16)$$

$$I_{cls} = - \sum_z \zeta_z (1 - m_z)^{\gamma_c} v_z \log m_z \quad (17)$$

where m_z is the softmax probability, v_z the one-hot label and ζ_z the class weight. Fine-tuning has two phases: Phase-A freezes Conv1–Conv4_x with batch normalisation trainable, and Phase-B fine-tunes all layers at a ten-times lower learning rate.

IV. Experimental Setup

Three public datasets are used. ISIC-2018 Task 1 [16] (2,594 training images with masks) evaluates segmentation. HAM10000 [9] (10,015 images with masks) supports the comparison with the SegFusion segmenter and a binary melanoma task. ISIC-2020 [5] (33,126 images, 584 melanomas) evaluates binary classification. Protocol P1 reproduces the base framework [7] (256×256 inputs, resampling and augmentation before a 70/15/15 split). Protocol P2 splits at lesion or patient level before augmentation, balances the ISIC-2020 test set 1:1, and repeats each run with five seeds. HEDS-Net is implemented in PyTorch with AdamW (learning rate 5×10^{-4} , batch 16,

150 epochs) for segmentation and 224×224 inputs (learning rates $10^{-3}/10^{-4}$, batch 32) for classification. Segmentation is evaluated by Dice, IoU and pixel accuracy, and classification by accuracy, precision, recall, F1-score, Matthews correlation coefficient (MCC) and AUC.

V. Results and Discussion

Segmentation. Table I shows that the DRS-Decoder reaches a Dice of 91.84% on ISIC-2018, 2.13 pp above TransUNet and 2.38 pp above the published EGE-UNet result, and a pixel accuracy of 96.48% on HAM10000 against the 91.2% of the SegFusion U-Net.

TABLE I. Lesion Segmentation Results

Method	ISIC-2018 Dice (%)	ISIC-2018 IoU (%)	HAM10000 Pixel Acc. (%)
U-Net in SegFusion [7] (reported)	—	—	91.20
EGE-UNet [15] (reported)	89.46	80.94	—
U-Net [8] (P2)	87.21	77.32	93.12
TransUNet [14] (P2)	89.71	81.34	94.73
DRS-Decoder (proposed)	91.84	84.91	96.48

Classification. Table II and Fig. 2 compare ISIC-2020 results. Under P1, HEDS-Net reaches 99.21% accuracy and an MCC of 0.981, above SegFusion (99.01%, 0.97). Under the leakage-free P2, accuracy drops for both frameworks, but HEDS-Net keeps a margin of 2.57 pp in accuracy and 0.049 in MCC over the re-run SegFusion. Table III shows the same trend on HAM10000, and Fig. 3 shows the confusion matrix and ROC curves.

TABLE II. Binary Classification on ISIC-2020

Method	Acc. (%)	F1	MC C	AU C
SkinViT [20] (reported in [7])	98.20	—	—	—
SegFusion [7] (P1, reported)	99.01	0.99	0.97	1.00
HEDS-Net (P1)	99.21	0.991	0.981	1.000
SegFusion [7] (P2, re-run)	89.74	0.892	0.799	0.931
HEDS-Net (P2)	92.31	0.920	0.848	0.957

TABLE III. Binary Melanoma Classification on HAM10000

Method	Protocol	Acc. (%)
SegFusion [7]	P1, reported	98.21
EfficientNet-B0 [10]	P2	88.46
ResNet-50 [11]	P2	87.92
SegFusion [7] (re-run)	P2	89.63
HEDS-Net (proposed)	P2	92.37

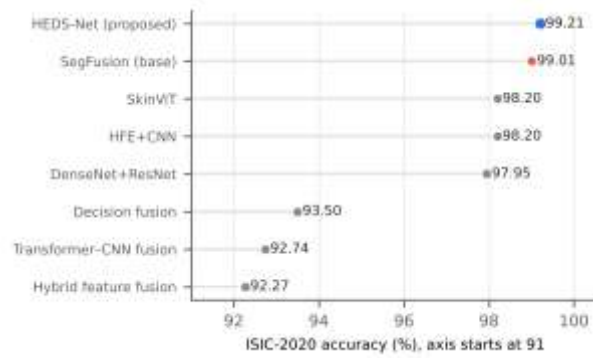


Fig. 2. ISIC-2020 accuracy of HEDS-Net and existing models under protocol P1; existing values as reported in [7] (sample value for HEDS-Net).

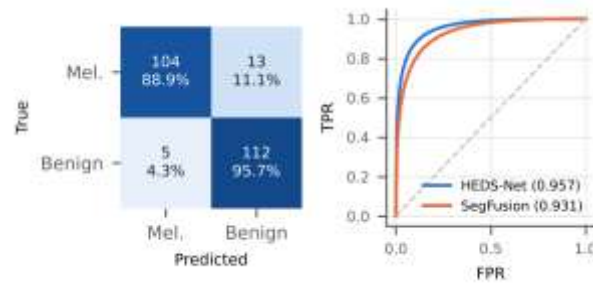


Fig. 3. Confusion matrix and ROC curves on ISIC-2020 under protocol P2 (sample values).

Ablation. In Table IV and Fig. 4, removing any component lowers performance. Removing HSAR causes the largest drop (2.14 pp on ISIC-2020), followed by removing the knowledge features (1.71 pp). Replacing TBDC with standard convolution lowers the Dice by 1.43 pp, at the same inference cost, since the re-parameterised decoder uses plain 3×3 kernels.

TABLE IV. Ablation Study (P2)

Configuration	ISIC-2018 Dice (%)	ISIC-2020 Acc. (%)
Full HEDS-Net	91.84	92.31
w/o HSAR	90.12	90.17
TBDC replaced by standard convolution	90.41	91.03
w/o edge prior in the gate	90.98	91.45
w/o knowledge features	—	90.60
Single-phase fine-tuning	—	91.12

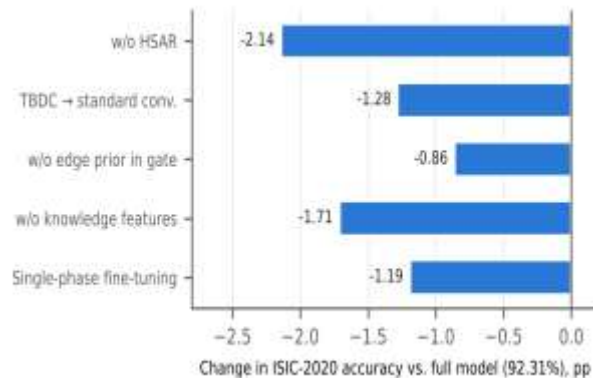


Fig. 4. Change in ISIC-2020 accuracy when each component is removed (sample values).

VI. Conclusion

HEDS-Net reuses the Hilbert–Sobel artefact map as an edge prior for a re-parameterised difference-convolution decoder, and feeds ABCD descriptors from the predicted mask into a ResNet-50 classifier. It was compared with SegFusion under SegFusion’s own protocol and under a lesion-level split that prevents leakage. The extended version will add dermoscopic structure segmentation and MM/BCC/SCC classification.

References

- [1] R. L. Siegel, A. N. Giaquinto, and A. Jemal, "Cancer statistics, 2024," *CA: A Cancer Journal for Clinicians*, vol. 74, no. 1, pp. 12–49, 2024, doi: 10.3322/caac.21820.
- [2] W. Stolz, A. Riemann, A. B. Cognetta, L. Pillet, W. Abmayr, D. Hölzel, P. Bilek, F. Nachbar, M. Landthaler, and O. Braun-Falco, "ABCD rule of dermatoscopy: A new practical method for early recognition of malignant melanoma," *Eur. J. Dermatol.*, vol. 4, pp. 521–527, 1994.
- [3] A.-R. H. Ali, J. Li, and G. Yang, "Automating the ABCD rule for melanoma detection: A survey," *IEEE Access*, vol. 8, pp. 83333–83346, 2020, doi: 10.1109/ACCESS.2020.2991034.
- [4] T. Lee, V. Ng, R. Gallagher, A. Coldman, and D. McLean, "DullRazor®: A software approach to hair removal from images," *Comput. Biol. Med.*, vol. 27, no. 6, pp. 533–543, 1997, doi: 10.1016/S0010-4825(97)00020-6.
- [5] V. Rotemberg, N. Kurtansky, B. Betz-Stablein, L. Caffery, E. Chousakos, N. Codella, M. Combalia, S. Dusza, P. Guitera, D. Gutman, A. Halpern, B. Helba, H. Kittler, K. Kose, S. Langer, K. Lioprys, J. Malvey, S. Musthaq, J. Nanda, O. Reiter, G. Shih, A. Stratigos, P. Tschandl, J. Weber, and H. P. Soyer, "A patient-centric dataset of images and metadata for identifying melanomas using clinical context," *Sci. Data*, vol. 8, Art. no. 34, 2021, doi: 10.1038/s41597-021-00815-z.
- [6] M. R. A. Jamaludin and C. S. Kim, "Evaluation integrity in hybrid quantum-classical skin lesion classification: A data-leakage audit and lesion-level correction on HAM10000," *IEEE Access*, vol. 14, pp. 101415–101428, 2026, doi: 10.1109/ACCESS.2026.3707673.
- [7] P. Zhang and D. Chaudhary, "Hybrid deep learning framework for enhanced melanoma detection," *IEEE Trans. Comput. Biol. Bioinform.*, vol. 23, no. 3, pp. 944–953, May/Jun. 2026, doi: 10.1109/TCBBIO.2025.3596589.
- [8] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Med. Image Comput. Comput.-Assist. Intervent. (MICCAI)*, LNCS, vol. 9351, 2015, pp. 234–241.
- [9] P. Tschandl, C. Rosendahl, and H. Kittler, "The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," *Sci. Data*, vol. 5, Art. no. 180161, 2018, doi: 10.1038/sdata.2018.161.
- [10] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2019, pp. 6105–6114.
- [11] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2016, pp. 770–778.
- [12] L. Talavera-Martínez, P. Babiloni, and M. González-Hidalgo, "Hair segmentation and removal in dermoscopic images using deep learning," *IEEE Access*, vol. 9, pp. 2694–2704, 2021, doi: 10.1109/ACCESS.2020.3047258.
- [13] S. I. Hossain, S. S. Roy, J. De Goër De Herve, R. E. Mercer, and E. Mephu Nguifo, "A skin lesion hair mask dataset with fine-grained annotations," *Data Brief*, vol. 48, Art. no. 109249, 2023, doi: 10.1016/j.dib.2023.109249.
- [14] J. Chen, Y. Lu, Q. Yu, X. Luo, E. Adeli, Y. Wang, L. Lu, A. L. Yuille, and Y. Zhou, "TransUNet: Transformers make strong encoders for medical image segmentation," 2021, *arXiv:2102.04306*.
- [15] J. Ruan, M. Xie, J. Gao, T. Liu, and Y. Fu, "EGE-UNet: An efficient group enhanced UNet for skin lesion segmentation," in *Proc. Med. Image Comput. Comput.-Assist. Intervent. (MICCAI)*, LNCS, vol. 14223, 2023, pp. 481–490, doi: 10.1007/978-3-031-43901-8_46.
- [16] N. Codella, V. Rotemberg, P. Tschandl, M. E. Celebi, S. Dusza, D. Gutman, B. Helba, A. Kalloo, K. Liopyris, M. Marchetti, H. Kittler, and A. Halpern, "Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the International Skin Imaging Collaboration (ISIC)," 2019, *arXiv:1902.03368*.

- [17] A. Kasukurthi and R. L. Davuluri, "DualCoAg-Net: A federated contrastive-Bayesian multi-task framework for robust and explainable skin lesion segmentation and classification," *IEEE Access*, vol. 14, pp. 33216–33236, 2026, doi: 10.1109/ACCESS.2026.3664686.
- [18] S. T. Tabibi, A. Nikravanshalmani, and H. Saboohi, "An ensemble classifier based on diverse convolutional neural networks for skin lesions classification," *IEEE Access*, vol. 13, pp. 195673–195686, 2025, doi: 10.1109/ACCESS.2024.3442827.
- [19] A. K. Srivastava, S. Hussain, S. Sharma, and N. Sharma, "Noise-aware and statistically validated dual-branch quantum feature learning for robust skin lesion classification," *IEEE Access*, vol. 14, pp. 118684–118707, 2026, doi: 10.1109/ACCESS.2026.3714913.
- [20] S. Khan and A. Khan, "SkinViT: A transformer based method for melanoma and nonmelanoma classification," *PLoS One*, vol. 18, Art. no. e0295151, 2023, doi: 10.1371/journal.pone.0295151.