UNIVERSITAS
GADJAH MADA
Kombinasi XGrad-CAM dan RefineCAM untuk Segmentasi Citra 3D MRI
Lathifah Aulia Nisa, Anugerah Galang Persada, S.T., M.Eng. ; Prof. Ir. Hanung Adi Nugroho, S.T., M.Eng., Ph.D., IPM
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

INTISARI

Segmentasi tumor otak berperan penting dalam mendukung diagnosis, perencanaan terapi, dan pemantauan penyakit. Model *deep learning*, khususnya 3D U-Net, banyak digunakan untuk tugas ini karena mampu mempelajari fitur secara otomatis dari data volumetrik. Namun, model tersebut bersifat *black box* sehingga sulit dipahami dasar pengambilan keputusannya. *Class Activation Mapping* (CAM) digunakan untuk mengatasi hal ini dengan menghasilkan *heatmap* area citra yang berkontribusi terhadap prediksi model. RefineCAM diusulkan untuk memperbaiki kualitas *heatmap* dengan menggabungkan atribusi dari beberapa lapisan jaringan, tetapi belum pernah diuji pada citra medis volumetrik maupun dikaji konsistensinya pada metode CAM yang berbeda. Penelitian ini bertujuan mengevaluasi performa model 3D U-Net dalam menyegmentasi tumor otak pada *dataset* BraTS2020, serta mengevaluasi pengaruh penerapan RefineCAM terhadap kualitas *heatmap* Grad-CAM sebagai metode dasar dan XGrad-CAM sebagai pengembangannya.

Model 3D U-Net dilatih menggunakan 235 pasien data latih dan divalidasi dengan 100 pasien, kemudian dievaluasi pada 34 pasien data uji menggunakan metrik Dice, *Jaccard*, *Recall*, dan *Specificity*. *Heatmap* Grad-CAM dan XGrad-CAM dihasilkan pada kondisi sebelum penerapan RefineCAM, yaitu dari satu lapisan *decoder* paling dangkal, dan sesudah penerapan RefineCAM, yaitu hasil penggabungan atribusi dari empat lapisan *decoder* melalui *MultiplierMix*. Kualitas kedua kondisi tersebut dibandingkan menggunakan metrik *Deletion* dan *Insertion Area Under the Curve* (DAUC dan IAUC), kemudian diuji signifikansinya secara statistik menggunakan uji Shapiro-Wilk yang dilanjutkan dengan *paired t-test* atau uji Wilcoxon *signed-rank*.

Hasil penelitian menunjukkan model 3D U-Net memperoleh Dice sebesar 0,9097 untuk *Whole Tumor*, 0,8666 untuk *Tumor Core*, dan 0,7347 untuk *Enhancing Tumor*. Pada evaluasi RefineCAM, XGrad-CAM menunjukkan perbaikan DAUC yang signifikan dari 0,0035 menjadi 0,0022, sedangkan Grad-CAM justru menunjukkan penurunan kualitas *heatmap* yang konsisten, dengan DAUC naik dari 0,0017 menjadi 0,0068. Temuan ini menunjukkan bahwa pengaruh RefineCAM tidak konsisten dan bergantung pada mekanisme pembobotan tiap metode CAM, sehingga penerapannya perlu mempertimbangkan karakteristik metode CAM yang digunakan.

Kata kunci : *segmentasi tumor otak, 3D U-Net, Class Activation Mapping, RefineCAM, explainable AI*

UNIVERSITAS
GADJAH MADA
Kombinasi XGrad-CAM dan RefineCAM untuk Segmentasi Citra 3D MRI
Lathifah Aulia Nisa, Anugerah Galang Persada, S.T., M.Eng. ; Prof. Ir. Hanung Adi Nugroho, S.T., M.Eng., Ph.D., IPM
Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/

ABSTRACT

Brain tumor segmentation plays an important role in supporting diagnosis, treatment planning, and disease monitoring. Deep learning models, particularly 3D U-Net, are widely used for this task because they can automatically learn features from volumetric data. However, such models remain black boxes, making it difficult to understand the basis of their predictions. Class Activation Mapping (CAM) addresses this limitation by producing heatmaps that highlight the image regions contributing to a model's prediction. RefineCAM was proposed to improve heatmap quality by combining attributions from multiple network layers, but it has never been tested on volumetric medical images, nor has its consistency across different CAM methods been examined. This study aims to evaluate the performance of a 3D U-Net model in segmenting brain tumors on the BraTS2020 dataset, as well as to evaluate the effect of applying RefineCAM on the heatmap quality of Grad-CAM as the base method and XGrad-CAM as its development.

The 3D U-Net model was trained on 235 patients and validated on 100 patients, then evaluated on 34 test patients using the Dice, Jaccard, Recall, and Specificity metrics. Grad-CAM and XGrad-CAM heatmaps were generated under two conditions, before applying RefineCAM using attribution from a single shallow decoder layer, and after applying RefineCAM using attribution combined from four decoder layers through MultiplierMix. The quality of both conditions was compared using the Deletion and Insertion Area Under the Curve (DAUC and IAUC) metrics, and the difference was tested for statistical significance using the Shapiro-Wilk test followed by a paired t-test or Wilcoxon signed-rank test.

The results show that the 3D U-Net model achieved a Dice score of 0.9097 for the Whole Tumor, 0.8666 for the Tumor Core, and 0.7347 for the Enhancing Tumor. In the RefineCAM evaluation, XGrad-CAM showed a significant improvement in DAUC from 0.0035 to 0.0022, whereas Grad-CAM instead showed a consistent decline in heatmap quality, with DAUC increasing from 0.0017 to 0.0068. These findings indicate that the effect of RefineCAM is not consistent and depends on the weighting mechanism of each CAM method, so its application needs to take into account the characteristics of the CAM method being used.

Keywords : *brain tumor segmentation, 3D U-Net, Class Activation Mapping, RefineCAM, explainable AI*