



UNIVERSITY  
OF WOLLONGONG  
AUSTRALIA

University of Wollongong  
Research Online

---

Faculty of Engineering and Information Sciences -  
Papers: Part B

Faculty of Engineering and Information Sciences

---

2017

# Assessing the prognostic impact of 3D CT image tumour rind texture features on lung cancer survival

Alanna Vial

*University of Wollongong, ahtv254@uowmail.edu.au*

David Stirling

*University of Wollongong, stirling@uow.edu.au*

Matthew Field

*Liverpool and Macarthur Cancer Therapy Centres, mfield@uow.edu.au*

Montserrat Ros

*University of Wollongong, montse@uow.edu.au*

Christian H. Ritz

*University of Wollongong, critz@uow.edu.au*

*See next page for additional authors*

---

## Publication Details

A. Vial, D. Stirling, M. Field, M. Ros, C. Ritz, M. Carolan, L. Holloway & A. A. Miller, "Assessing the prognostic impact of 3D CT image tumour rind texture features on lung cancer survival," in in IEEE Global Conference on Signal and Information Processing (GlobalSIP 2017), 2017, pp. 735-739.

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library:  
research-pubs@uow.edu.au

---

# Assessing the prognostic impact of 3D CT image tumour rind texture features on lung cancer survival

## **Abstract**

In this paper we examine a technique for developing prognostic image characteristics, termed radiomics, for non-small cell lung cancer based on a tumour edge region-based analysis. Texture features were extracted from the rind of the tumour in a publicly available 3D CT data set to predict two-year survival. The derived models were compared against the previous methods of training radiomic signatures that are descriptive of the whole tumour volume. Radiomic features derived solely from regions external, but neighbouring, the tumour were shown to also have prognostic value. By using additional texture features an increase in accuracy, of 3%, is shown over previous approaches for predicting two-year survival, upon examining the outside rind including the volume compared to the volume without the rind. This indicates that while the centre of the tumour is currently the main clinical target for radiotherapy treatment, the tissue immediately around the tumour is also clinically important.

## **Disciplines**

Engineering | Science and Technology Studies

## **Publication Details**

A. Vial, D. Stirling, M. Field, M. Ros, C. Ritz, M. Carolan, L. Holloway & A. A. Miller, "Assessing the prognostic impact of 3D CT image tumour rind texture features on lung cancer survival," in in IEEE Global Conference on Signal and Information Processing (GlobalSIP 2017), 2017, pp. 735-739.

## **Authors**

Alanna Vial, David Stirling, Matthew Field, Montserrat Ros, Christian H. Ritz, Martin G. Carolan, Lois C. Holloway, and Andrew Alexis Miller

# ASSESSING THE PROGNOSTIC IMPACT OF 3D CT IMAGE TUMOUR RIND TEXTURE FEATURES ON LUNG CANCER SURVIVAL MODELLING

Alanna Vial<sup>1</sup>  
Christian Ritz<sup>1</sup>

David Stirling<sup>1</sup>  
Martin Carolan<sup>1,2</sup>

Matthew Field<sup>3</sup>  
Lois Holloway<sup>1,3</sup>

Montserrat Ros<sup>1</sup>  
Alexis A. Miller<sup>1,2</sup>

<sup>1</sup> School of Electrical, Telecommunications and Computer Engineering,  
University of Wollongong, Wollongong, NSW, AUSTRALIA

<sup>2</sup> Illawarra Cancer Care Centre, Wollongong, NSW, AUSTRALIA

<sup>3</sup> Liverpool and Macarthur Cancer Therapy Centres and Ingham Institute, Liverpool, NSW, AUSTRALIA

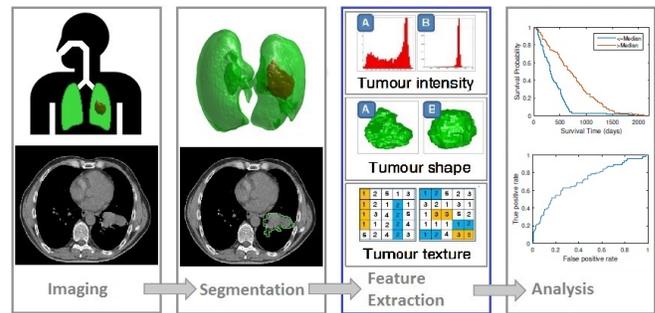
## ABSTRACT

In this paper we examine a technique for developing prognostic image characteristics, termed radiomics, for non-small cell lung cancer based on a tumour edge region-based analysis. Texture features were extracted from the rind of the tumour in a publicly available 3D CT data set to predict two-year survival. The derived models were compared against the previous methods of training radiomic signatures that are descriptive of the whole tumour volume. Radiomic features derived solely from regions external, but neighbouring, the tumour were shown to also have prognostic value. By using additional texture features an increase in accuracy, of 3%, is shown over previous approaches for predicting two-year survival, upon examining the outside rind including the volume compared to the volume without the rind. This indicates that while the centre of the tumour is currently the main clinical target for radiotherapy treatment, the tissue immediately around the tumour is also clinically important.

**Index Terms**— Image Processing, Machine Learning, Non-Small Cell Lung Cancer, Radiomics

## 1. INTRODUCTION

Radiomics is an analytic approach to extract quantitative features, such as shape, texture or intensity, within radiological imaging data, such as computed tomography (CT), positron emission tomography (PET) and magnetic resonance imaging (MRI), by utilising existing image processing techniques. The information from radiomic analyses can be correlated with patient data, such as oncologist-defined regions of interest (ROI), outcome (tumour control, survival), and tumour phenotype represented in histopathology, or biomarkers. Radiomic information could inform the selection of cancer therapy for an individual patient and tumour, and could predict clinical outcomes. The four main stages of radiomic analysis include imaging, segmentation, feature extraction and analysis for classification and prediction of an outcome such as two-year survival, see Figure 1.



**Fig. 1:** This figure conveys the process of using image derived features (radiomics) for prognostic models. The third segment was obtained from [1].

In this study we explore the textural radiomic features of a non-small cell lung cancer (NSCLC) tumour defined by an oncologist. The marked tumour volume is denoted as the Gross Tumour Volume (GTV).

The GTV defined by the oncologist produces a whole 3D volume. Taking this volume, we then defined a layer of known thickness of tissue outside the GTV-line by adjoining the GTV (outer rind), and a corresponding layer of identical thickness inside and abutting the GTV-line (inner rind) [2]. These volumes were then used for the analysis of radiomic signatures and two-year survival outcome. Figure 2 illustrates the inside and outside rinds of the tumour volume, where green indicates the outside rind and blue the inside rind, furthermore red shows the gross tumour volume (GTV), which has been excluded from the analysis.

We explored the effect on predicting two-year survival by including only imaging textures discussed in Aerts et al. [1]. We used the same public NSCLC dataset that was utilised in [1] to predict the two-year survival. To the best of our knowledge this is the first study exploring the rind of a tumour using radiomic texture analysis. The rationale for observing the rind region is to define a strategy for partitioning and analysing the texture of tumour regions so as to characterise the

relationship between texture and survival in NSCLC patients.

## 2. RELATED WORK

An earlier report demonstrated that four radiomics features could predict the two-year survival of lung cancer patients in a public NSCLC dataset with a high degree of accuracy, in Aerts et al. [1]. We have undertaken a similar analysis using three of these identified highly prognostic features to understand whether the presence of these radiomic features in the inner and outer rinds around the GTV-line can also predict two-year survival. Due to unforeseen technical issues the fourth wavelet feature was not able to be included in this study. The construction and use of these rinds in radiomics analysis is unique.

Subramaniam et al. [3] extended the previous work in NSCLC by quantifying the heterogeneity of  $SUV_{max}$  values from a PET scan. By using Kaplan Meier survival analysis combined with Cox Proportional Hazard regression, they were able to identify ‘good’ and ‘poor’ prognosis groups with an improvement in survival prediction.

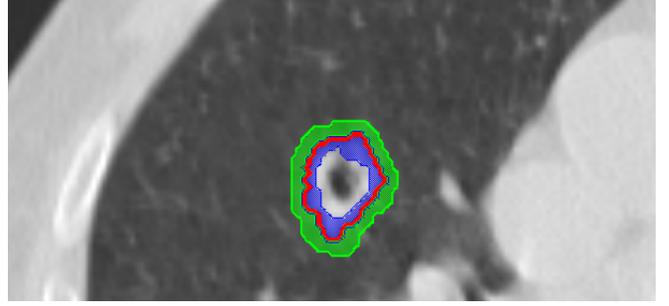
Tumour heterogeneity for NSCLC patients was further studied by Ganeshan et al. [4], using the technique computer tomography texture analysis (CTTA), that allows coarse and fine image texture features to be filtered, showing a high correlation with survival. Textural analysis has been utilised for classification of medical outcomes in the past [5], [6].

Given that NSCLC commonly develops distant metastases (DM) leading to patient mortality, Coroller et al. [7] sought radiomic signatures which were prognostic for DM in NSCLC patients, again improving survival prediction. They found that Gray-Level Co-occurrence Matrix texture features were useful. Vallières et al. [8] undertook a similar study examining a GTV defined on pretreatment MRI fused with a PET scan in soft tissue sarcomas (STS) to evaluate the risk of DM in the lung. They showed that the fused MRI/PET image set had a superior performance in predicting DM.

These related works reveal the potential for radiomics to improve the accuracy of outcome predictions. Until now, reported analyses have focused on the tumour (GTV) features alone. We have begun to analyse the lung exterior and surrounding the tumour also. This approach is supported by the approach of radiation oncologists who encapsulate the visible GTV within a Clinical Target Volume (CTV) which defines the at risk areas which appear to be normal to the naked eye.

## 3. PROPOSED METHODOLOGY

This analysis consists of four main stages; converting the DICOM-RT files into matrices readable in MATLAB, calculating the rind volume for extraction from the CT data, determining the radiomic features and finally producing the two-year survival analysis.



**Fig. 2:** Illustration of the rind concept shown on a single CT slice of a patient NSCLC image set.

The DICOM-RT format is the standard medical imaging format used to store medical imaging data and is paramount in the efficacious application of the picture archiving and communications systems (PACS) used in most hospitals for radiation therapy [9]. It is essentially a file format that allows multiple CT images to be related using an overarching RT-STRUCT file which describes the oncology delineation data and the dosage data for the patient.

422 NSCLC patient data sets were publicly available at the time of publication, 245 were usable in this study, approximately 14.2GB from over 29,000 CT image files were examined, with the files obtained from [1], [10], [11].

To derive the mask, we first define the set of points representing the inside and outside rind, let  $\mathbf{x} \in \mathbb{R}^3$  and  $\mathbf{z} \in \mathbb{R}^3$  define positions vectors in the 3-Dimensional image set frame. Further, let  $\mathcal{V} \subset \mathbb{R}^3$  be the set of all points in the gross tumour volume (GTV) as given by the expert or any other segmented region of interest (ROI) and  $\hat{\mathcal{V}} \subset \mathcal{V}$  is defined as all points on the boundary of this region.

Suppose there is a distance function that can be constructed for each ROI,  $d(\hat{\mathcal{V}}) : \mathbb{R}^3 \rightarrow \mathbb{R}$ , which will be defined by the following

$$d(\mathbf{x}, \hat{\mathcal{V}}) = \begin{cases} \min \|\mathbf{x} - \mathbf{z}\|, & \forall \mathbf{z} \in \hat{\mathcal{V}}, \forall \mathbf{x} \notin \mathcal{V} \\ -\min \|\mathbf{x} - \mathbf{z}\|, & \forall \mathbf{z} \in \hat{\mathcal{V}}, \forall \mathbf{x} \in \mathcal{V} \end{cases}$$

Now the outer boundary region can be expressed by the set

$$S_o = \{\mathbf{x} \mid 0 < d(\mathbf{x}, \hat{\mathcal{V}}) < L\}, \quad (1)$$

where  $L$  is the width of a region that contains all line segments perpendicular to the surface.  $S_o$  is therefore a region external to the ROI. Conversely, we have

$$S_i = \{\mathbf{x} \mid -L < d(\mathbf{x}, \hat{\mathcal{V}}) < 0\}, \quad (2)$$

defining an inner region such that  $S_i \subseteq \mathcal{V}$ . Colloquially, we refer to the region,  $S = \{S_i \cup S_o\}$ , as the rind of a given ROI, obtained by expanding or contracting the ROI uniformly from the surface by  $L$ . This can be visualised in Figure 2 where the 3D rind is found by first finding the 2D rind for each image.

This rind was then dilated and contracted to ensure there were no pixels missing within the rind, creating a rind exactly  $2 \times n$ , being 6 pixels deep in this study, all the way around the 3D GTV, this is an expansion of approximately 6mm around the whole GTV volume. This can be seen in Figure 3.

We then extracted the three features recognised by the authors in [1] as having prognostic significance and compared the two-year survival predictions for the whole volume compared to the inside rind and outside rind alone. These features included the statistical energy, shape compactness and the run length non-uniformity of the image, these features are used for comparison in this study.

Let  $\mathbf{X}$  signify the 3-Dimensional image matrix, containing  $N$  voxels, energy is given by equation 3.

$$energy = \sum_i^N \mathbf{X}(i)^2 \quad (3)$$

Let  $V$  represent the volume of the tumour, while  $A$  is the surface area of the tumour, shape compactness is given by equation 4.

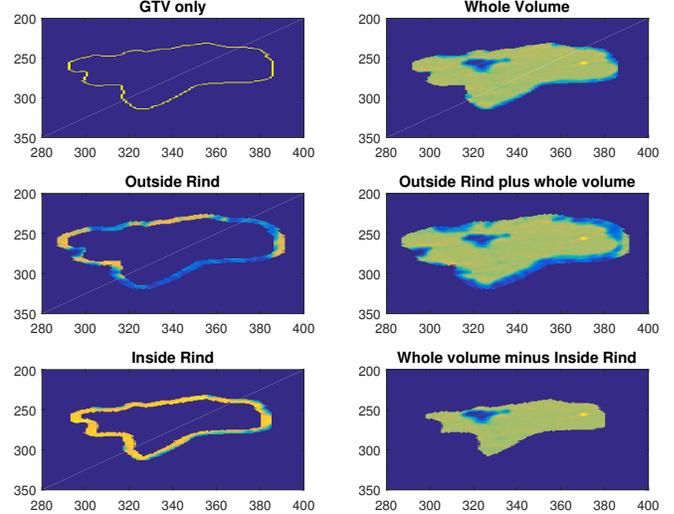
$$shape = 36\pi \frac{V^2}{A^3} \quad (4)$$

Let  $p(i, j|\theta)$  be the  $(i, j)$ th entry in the given run-length matrix  $p$  for a direction  $\theta$ . Let  $N_g$  the number of discrete intensity values in the image and  $N_r$  be the number of different run lengths. The run length non-uniformity (RLN) is given by equation

$$RLN = \frac{\sum_{j=1}^{N_r} \left[ \sum_{i=1}^{N_g} p(i, j|\theta) \right]^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i, j|\theta)} \quad (5)$$

The second study inspected the 52 texture features described in Aerts et al. [1] and Vallières et al. [8], these features look at the spatial distribution of voxel intensities using gray level run-length and gray level co-occurrence matrices. In Aerts et al. [1] four features were selected from 440, in this study we explore the prognostic significance of the three features described above in equations 3,4 and 5, excluding the wavelet feature.

The data, consisted of DICOM-RT CT image files along with two-year survival outcome for 245 NSCLC patients. A 10-fold cross-validation approach was employed to evaluate the models. Each fold of the cross-validation was stratified based on the propensity of outcomes in the total cohort. Logistic regression was then employed to produce the two-year survival Kaplan-Meier curves, as this is very similar to the multi-variant cox proportional hazards regression model used by Aerts et al. [1]. These curves are often used in medical research to predict survival by stratifying survival likelihood into two different patient groups [1], [3], [4]. The data was then separated into a good and bad prognosis group separated by the median values.



**Fig. 3:** Example outside rind compared to inside rind and whole tumour volume for a single CT slice.

## 4. RESULTS

The results can be seen in Figure 4. As a similar analysis we looked at the 52 texture based features, which showed an improvement in the two-year survival, as shown in Figure 5.

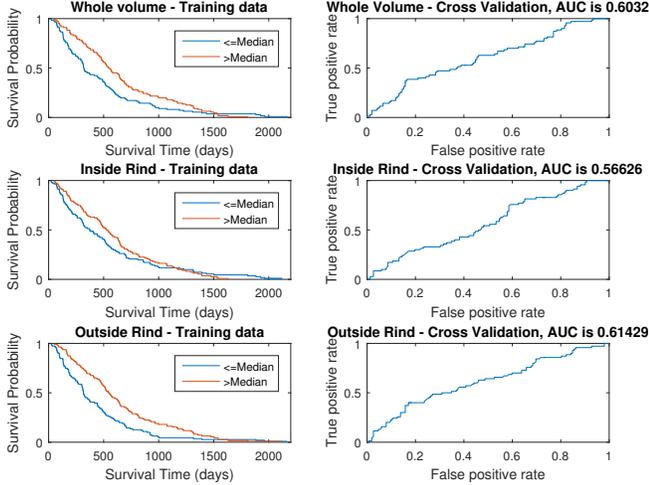
It was found that the outside rind could predict the survival better than the whole volume alone. As a result of this discovery, we included an evaluation of the outer rind combined with the whole volume in this study.

Analysing the results for all texture features, we found the outside rind plus the whole volume to produce the best prediction for two-year survival, this was closely followed by the inside rind and the whole volume. The whole tumour volume without the inside rind produced the worst results, when comparing the area under the curve (AUC) values after twenty iterations of the cross-validation, shown in Table 1.

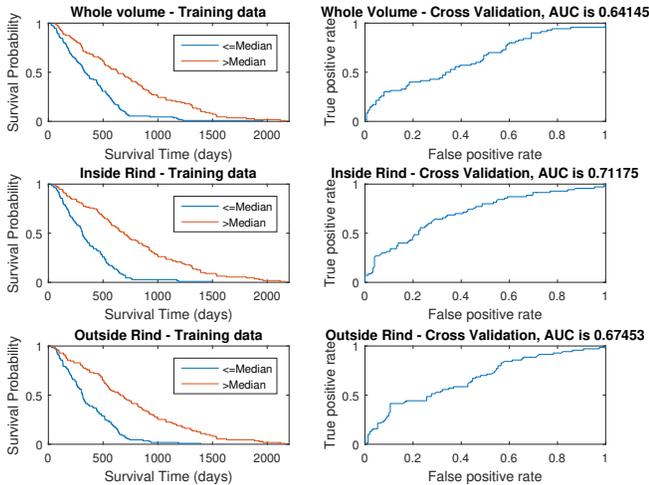
Overall results of all textures features was approximately 10% better at predicting survival than the three feature analysis alone. The Kaplan Meier curves in Figure 5 also demonstrate that the models based on all texture features improve in discriminating survivors from non-survivors. The bottom blue curve represents the poor prognosis group and the top red curve represents the good prognosis group. Increased separation between these curves indicates greater success in two-year survival classification. This is clearly demonstrated in the case of all texture features being tested however in the case of the three textures only, the difference between these two curves is less significant, see Figure 4.

## 5. DISCUSSION

This work presents analytic approaches to delineate a tumour and describes features that produce the best two-year survival



**Fig. 4:** The diagrams on the left show the logistic regression of energy, shape and grey level non-uniformity radiomic signatures, while the models on the right show the ROC for classification of the whole volume, inside rind and outside rind for the training data compared to the validation data.



**Fig. 5:** The graphs on the left show the logistic regression of all 52 texture based radiomic features, while the models on the right show the ROC for classification of the whole volume, inside rind and outside rind for the training data compared to the validation data.

prediction. The clinical perspective is that the most important region of the tumour to target is the centre of the tumour, however these results found the whole tumour without the rind (just the centre of the tumour) has the worst performance. In clinical practice, a volume outside the GTV called the clinical target volume (CTV) is added to try to include cancer infiltrating into surrounding lung that cannot be appreciated by the naked eye.

This work supports this clinical practice by showing that

**Table 1:** This table provides the AUC results from 20 iterations for all texture features compared to the three feature analysis, in addition to the standard deviation.

Volume Analysed	Three Features	Texture Features
	AUC $\pm\sigma$	AUC $\pm\sigma$
Whole vol.	0.589 $\pm$ 0.015	0.661 $\pm$ 0.016
Inner rind	0.558 $\pm$ 0.011	0.679 $\pm$ 0.024
Outer rind	<b>0.598<math>\pm</math>0.011</b>	0.689 $\pm$ 0.015
Outer rind with vol.	0.583 $\pm$ 0.013	<b>0.699<math>\pm</math> 0.011</b>
Vol. excl. inner rind	0.584 $\pm$ 0.013	0.624 $\pm$ 0.015

important textual information that relates to two-year survival is not necessarily within the tumour but around it. The likely mechanism for this improvement in predictive performance arises from the ability of radiomic analysis to find pertinent signals within the image that cannot be appreciated by the eyes of an oncologist, thereby determining the extent of the tumour more accurately. The post hoc analysis of treatment images will identify visually imperceptible instances of 'geographic miss' which will result in worse survival.

## 6. CONCLUSION AND FUTURE WORK

In this paper we explored the textual radiomic features in a whole 3D tumour volume, compared to the inside and outside rind of the tumour only, for CT images of NSCLC. The textual features found within the tumour rinds were very similar to the textures in the entire tumour volume, and this analysis predicted two-year survival with an improved accuracy of 3% for survival classification using textures from the outside rind compared to the whole volume. It is important to note, that there is significant clinical uncertainty in defining the tumour boundaries of the GTV and that this should be considered when viewing these results.

One limitation of this work is in the manual selection of the radiomic features. In future work we will use a machine learning algorithm such as convolutional neural networks [12] to determine which radiomic features have the highest prognostic significance for determining two-year survival. In addition, the radiomic features themselves are determined based on general mathematical formulas associated with analysing an image, in a future study we plan to use modern image processing techniques to determine alternative radiomic features.

## 7. ACKNOWLEDGEMENTS

This research has been conducted with the support of the Australian Government Research Training Program Scholarship, along with the Radiation Oncology Staff Specialist Trust Funds at Illawarra Cancer Care Centre, Wollongong NSW and the Liverpool Cancer Therapy Centre, Liverpool NSW.

## References

- [1] H. J. W. L. Aerts et al., “Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach,” *Nature Communications*, vol. 5, 2014.
- [2] C. McGibney et al., “Analysis of dose distribution in the ‘Rind’-a volume outside the PTV-in 3-dimensional conformal radiation therapy of non-small cell lung cancer,” *Radiotherapy and Oncology*, vol. 66, no. 1, pp. 87–93, May 2003.
- [3] A. R. Subramaniam et al., “A novel metric for quantification of homogeneous and heterogeneous tumors in PET for enhanced clinical outcome prediction,” *Physics in Medicine and Biology*, vol. 61, no. 1, p. 227, 2016.
- [4] B. Ganeshan et al., “Tumour heterogeneity in non-small cell lung carcinoma assessed by CT texture analysis: a potential marker of survival,” *European Radiology*, vol. 22, no. 4, pp. 796–802, 2012.
- [5] Rahmadwati and G. Naghdy et al., “Classification cervical cancer using histology images,” in *2010 2nd International Conference on Computer Engineering and Applications, ICCEA 2010*, vol. 1, 2010, pp. 515–519.
- [6] T. A. Alshawi, F. E. A. El-Samie, and S. A. Alshebeili, “Magnetic resonance and computed tomography image fusion using bidimensional empirical mode decomposition,” in *2015 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Dec. 2015, pp. 413–417.
- [7] T. P. Coroller et al., “CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma,” *Radiotherapy and Oncology*, vol. 114, no. 3, pp. 345–350, 2015.
- [8] M. Vallières et al., “A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities,” *Physics in Medicine and Biology*, vol. 60, no. 14, p. 5471, 2015.
- [9] M. Y. Y. Law and B. Liu, “DICOM-RT and Its Utilization in Radiation Therapy,” *RadioGraphics*, vol. 29, no. 3, pp. 655–667, May 2009.
- [10] K. Clark et al., “The cancer imaging archive (TCIA): Maintaining and operating a public information repository,” *Journal of Digital Imaging*, vol. 26, no. 6, 2013.
- [11] H. J. W. L. Aerts et al. (2015). Data From NSCLC-Radiomics. The Cancer Imaging Archive. <http://doi.org/10.7937/K9/TCIA.2015.PF0M9REI>.
- [12] Y. LeCun et al., “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.