B.TECH. PROJECT REPORT

ON

High-Fidelity Damage Classification of milled rice grains using Deep Unsupervised Learning

BY

Kriz Deryll Moses

DISCIPLINE OF MECHANICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY INDORE

May 2022

High-Fidelity Damage Classification of milled rice grains using Deep Unsupervised Learning

A PROJECT REPORT

Submitted in partial fulfilment of the

requirements for the award of the degrees

of

BACHELOR OF TECHNOLOGY

In

MECHANICAL ENGINEERING

Submitted by:
Kriz Deryll Moses

Guided by:

Dr. Pavan Kumar Kankar

INDIAN INSTITUTE OF TECHNOLOGY INDORE
May 2022

CANDIDATE'S DECLARATION

I hereby declare that the project entitled "**High-Fidelity Damage Classification of milled rice grains using Deep Unsupervised Learning**" submitted in partial fulfilment for the award of the degree of Bachelor of Technology in Mechanical Engineering completed under the supervision of **Dr Pavan Kumar Kankar, Associate Professor, IIT Indore** is an authentic work.

       Further, I declare that I have not submitted this work for the award of any other degree elsewhere.

<div align="right">

Name: Kriz Deryll Moses
Date: 25/05/2022

</div>

Signature and name of the student(s) with date

_____

CERTIFICATE by BTP Guide(s)

       It is certified that the above statement made by the student(s) is correct to the best of my/our knowledge.

25/05/2022   Dr. Pavan K. Kankar

Associate Professor, ME

Signature of BTP Guide(s) with dates and their designation

Preface

This report on "High-Fidelity Damage Classification of milled rice grains using Deep Unsupervised Learning" is prepared under the guidance of Dr Pavan Kumar Kankar, Associate Professor, IIT Indore.

Through this report, I explained the development of a deep unsupervised method *Contrastive-RC* for fine-grained damage classification of milled rice grains, leveraging self-supervised contrastive learning technique. I have tried to explain every aspect of the method, covering the database used, explaining the model architecture, along with providing an in-depth analysis of its performance on the test images. The method Contrastive-RC not only facilitates the broader classification of rice grain images into six damage-based classes with well-defined features, but also provides a means for further subclassification providing low level control allowing for it to be used in multiple use cases

I have tried to the best of my ability and knowledge to explain the content in a simplified manner by using illustrative diagrams and a step by step approach.

Kriz Deryll Moses

B.Tech. IV Year

Discipline of Mechanical Engineering

IIT Indore

Acknowledgements

I wish to thank Dr. Pavan Kumar Kankar, Associate Professor, for his kind support and valuable guidance.

It is with his help and support that I was able to complete the design and technical report.

Without his support, this report would not have been possible.

Kriz Deryll Moses

B.Tech. IV Year

Discipline of Mechanical Engineering

IIT Indore

Abstract

The quality evaluation of processed rice grains is an important factor in determining market acceptance, pricing, storage stability, processing quality, and overall consumer approval. Damage classification based on visual symptoms of raw rice grains allows for very effective quality evaluation. In the current literature, the machine vision methods are predominantly based on supervised machine learning which is fundamentally dependent on manual labelling. However, manual labelling faces issues like erroneousness, subjectiveness and overlapping classes. There exists no work in the current literature which presents an unsupervised approach to classifying rice grain damages. In this study, a deep unsupervised method Contrastive-RC [KM1] is developed for fine-grained damage classification of processed white rice, leveraging contrastive self-supervised learning technique. In particular, self-supervised contrastive learning (SimCLR) is used for feature representation followed by dimensionality reduction (UMAP) and clustering (HDBSCAN). For this, a large dataset of 20,134 high magnification (24 MP) images of individual rice grains spread across different damages was collected. I have been successful in clustering the rice grains into six main cluster based classes with well-defined attributes. The class names along with the number of corresponding instances are: *normal-damage* (5599), *chalky-discoloured* (4987), *discoloured* (3215), *half-chalky* (2386), *healthy* (2061), and *broken* (931). Further, it is also presented how the method can be extended to subclassify these damages according to the user's needs by providing a low-level control, enabling the method to be used in multiple use-cases. The method is fast, versatile and robust towards changes in messy variables like brightness, grain orientation, etc., making it ideal for real world use and extension to other varieties of white processed rice. Overall, this study presents a deep unsupervised method Contrastive-RC for fine-grained damage classification of processed white rice, leveraging contrastive self-supervised learning technique, which could be utilised as a tool for better and more objective quality assessment of the damaged rice grains at market and trading locations.

Table of Contents

List of Figures

Chapter 1 - Introduction

1.1 Background

Rice is one of the most important food grains in the world. It is India's largest produced food grain crop, cultivated on approximately 34% of total cropped land in fiscal year 2021, with its production accounting for 42% of the country's fresh produce crop production [4]. Furthermore, India is the world's second-largest producer and the largest exporter of rice. In the recent decades, there has been a steady increase in the need for high-quality food due to increasing demand, knowledgeable consumers and factors like health and environmental concerns [2]. As a result, rice quality evaluation has become increasingly important. The quality evaluation of processed (or milled) rice grains is an important factor in market acceptability, pricing, storage stability, processing quality, and overall consumer approval [5].

1.2 Damage based quality evaluation of rice

The entire process of rice quality evaluation is complex and requires a lot of domain expertise. Its complexity can broadly be broken down into two parts: 1) The criterion of evaluation: Setting well defined attributes which can be used to mark the quality; 2) The method used: An ideal method would be able to evaluate the quality in a timely, cost-effective, non-destructive and accurate manner. In India, quality evaluation is mainly performed manually by classifying grains into different categories and marking the relative quality for each.

Classification of grains into various categories allows for very effective quality evaluation. In general, this classification can be based on variety and/or damage. Variety-based classification categorises distinct rice varieties based on very generic factors such as size, shape, taste, aroma, kernel hardness, moisture content, or visual features such as the presence of weakened, infected, discoloured, or foreign objects [6]. Damage-based classification, on the other hand, differentiates grains based on the type of damage present in them. That is not only identifying the presence of a damage but also determining its type.

In this study, the focus is on the damage classification of rice. There are a number of studies on rice grain damages including [7-11]. It is seen that the visual features of raw rice grains can

be very effectively used in damage identification and hence become key quality factors [5, 6]. The USDA [79] attempts to provide visual references for the common damages found in rice. Some of the common features used to define damages in rice grains include broken kernel, chalkiness, discoloration, husk layer, surface indentations, spots (watermark, insect damage, stain) and highly dark regions signifying heat damage. It is common for grains to occur with multiple symptoms. These are demonstrated in Fig. 1 and can have a significant impact on the market value of rice [89, 98, 6].
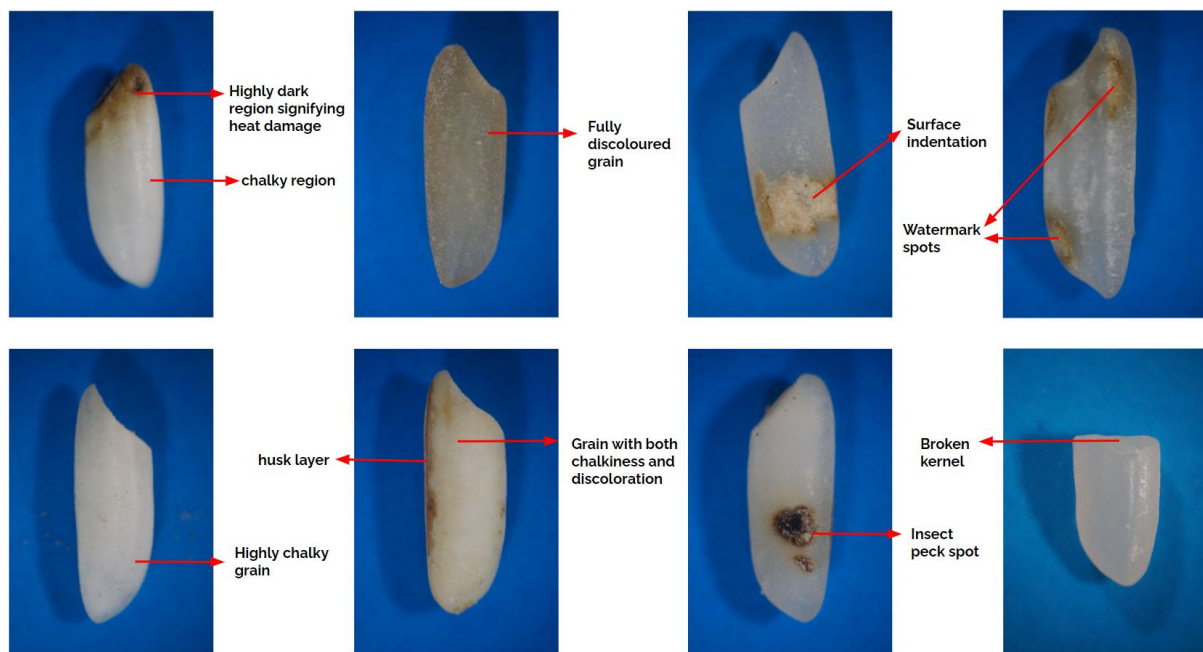
Figure 1: Rice grains with different types of damages. Commonly used visual symptoms to identify damages in processed rice are broken kernel, chalkiness, discoloration, husk layer, surface indentations, spots (watermark, insect damage, stain) and highly dark regions signifying heat damage. It is common for grains to contain multiple damages.

1.3 Quality evaluation of rice in India

In India, the common practices to evaluate rice quality work on the basis of the number of damages present in them. Instead of separating rice into different damages, there is a limit set on the percentage of each damage present in the rice stock. When rough rice is milled, it yields both head rice and a variety of by-products, including brokens, brewers, bran, mill feed, and hulls. In terms of revenue, head rice is the most significant. It refers to the grains with length more than 75% of that of an unbroken grain. Red rice, weed seeds, damaged kernels, off-colour,

chalky, and other sorts of rice kernels are among the quality parameters that grades aim to quantify.

According to the official inspection criteria (CNS, 1995) [12], the top grade of rice must contain over 75 % sound kernels, not more than 4 % dead kernels, not more than 3 % chalky kernels, and not more than 3.7 % total of damaged, broken, off–type kernels, and paddy (rough rice). As per circular issued by FCI, Category 'A' (upto 3% damaged grain and upto 3% discoloured grain), Category 'B' (above 3% upto 4% damaged grain and above 3% upto 5% discoloured grain), and Category 'C' (above 4% upto 5% damaged grain and above 5% upto 7% discoloured grain) are all fit for human consumption. Rice stocks with more than 5% damaged grains must be regarded as Non-Issuable Stock, which is unsuitable for human consumption. These stocks further categorised as: 1. Fit for cattle feed, 2. Fit for poultry feed, 3. Fit for industrial use, 4. Fit for manure, 5. Fit for dumping [14]. Damaged grains that are broken, cracked, attacked by fungi or insects, etc. without nutritional value, still could be used as suitable alternative sources for economically and environmentally bioethanol production [16]. Also, the Government of India has launched Ethanol Blended Petrol Programme (EBP Programme) and has scaled up blending targets from 5% to 10%.

Since the quality as well as the price of a batch of rice is dependent on the percentage of damaged grains present in it, and different damage types have different values, it becomes critical to identify each damage type with exact numbers. In the literature, the word "damage" is used with different meanings. Some refer to chalky, discoloured and broken grains separately from damaged grains. Others include them in the damaged grains category. In this study, we present six damage based classes in processed white rice: *healthy*, *broken*, *half-chalky*, *chalky-discoloured*, *discoloured* and *normal-damage*. All classes apart from the healthy class are considered to be damage classes. We refer to the term "damaged" as being inferior from the healthy grain; having some kind of defect. This inferiority is shown in terms of market value as the grains in the five damage classes all have lesser market values compared to that of grains in the healthy class.

1.4 Research Gap

Damage-based classification would typically be performed to classify different grains of a single variety. This can be used to identify the different types of damages present in any rice stock, and to further segregate them according to the various market needs. However, in India, rice quality is mostly measured manually by experienced personnel which is time consuming, expensive and error-prone. Lab tests are a reliable alternative but it is more expensive, time-consuming, and difficult to set.

As quality inspection based on damage classification can be performed very effectively using visual symptoms, Machine Vision (MV) becomes an ideal approach to its solution. Machine vision provides an automated, non-destructive, and cost-effective way of determining the grain quality based on the visual symptoms [5, 19, 20]. The industry standard state-of-the-art Sortex machines [18] which are used for rice segregation are based on machine vision systems. Although, they are extremely expensive and have limited functionality, in that they sort the grains into healthy versus damaged without quantifying the type and the degree of damage, making them unsuitable for further grain separation. Other very commonly used systems include Flatbed scanners [17], which are highly dependent on the algorithm used for classification.

Recently, computer vision methods based on machine learning [21] and in particular deep learning algorithms have become very popular in the field of agriculture [22, 23], and have found application in rice categorisation as well. There has been a considerable development in the classification of rice using machine vision-based systems as reviewed in Section 2.1 in detail. However, the methods in the existing literature have been constrained in one way or another in their approaches. Most of the methods have focused on variety classification rather than damage classification [28-30, 36-40, 47]. Some have attempted to classify rice into different grades based on shape and size alone [25, 26]. The machine learning methods used have been mostly supervised and the labelled classes have been fairly trivial to separate, with well-defined boundaries between them easily evident to the human eye [28, 29, 30]. Same was observed in [31] and [41] which attempted damage classification into (cracked, chalky, damaged and spotted) and (broken, chalky, red-spotted and black-spotted) classes, respectively. Since there is a clear separation between such classes, the features used to distinguish are much easier to extract and hence the classification can be performed with smaller datasets (< 500 samples per class) and lower resolution images. Further, manual

labelling is straightforward in such cases. However, in this study a fine-grained damage classification is attempted, in which features are complex and there exists the problem of overlapping classes as demonstrated in this study. This poses a more difficult problem and requires a large size representative dataset which captures all the different types of grains you can find in processed rice. In this study, we prepared a dataset of 20,134 images of single kernels of white processed rice spread across various damages.
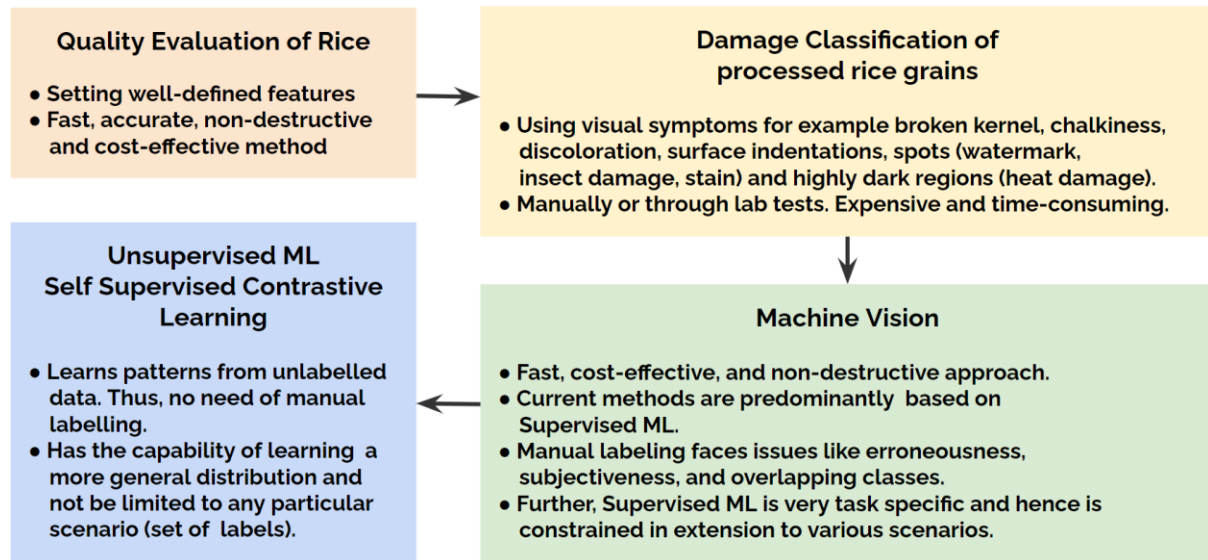
Quality Evaluation of Rice
- Setting well-defined features
- Fast, accurate, non-destructive and cost-effective method

Damage Classification of processed rice grains
- Using visual symptoms for example broken kernel, chalkiness, discoloration, surface indentations, spots (watermark, insect damage, stain) and highly dark regions (heat damage).
- Manually or through lab tests. Expensive and time-consuming.

Unsupervised ML Self Supervised Contrastive Learning
- Learns patterns from unlabelled data. Thus, no need of manual labelling.
- Has the capability of learning a more general distribution and not be limited to any particular scenario (set of labels).

Machine Vision
- Fast, cost-effective, and non-destructive approach.
- Current methods are predominantly based on Supervised ML.
- Manual labeling faces issues like erroneousness, subjectiveness, and overlapping classes.
- Further, Supervised ML is very task specific and hence is constrained in extension to various scenarios.

Figure 2. Problem Background

The current methods in the literature are predominantly based on supervised machine learning, which is fundamentally dependent on the labelling provided. However, in damage classification of rice, manual labelling as discussed, faces issues like subjectiveness, erroneousness and overlapping damages. Further supervised machine learning is task specific, which acts as a constraint when a model on one problem needs to be extended to another where a different set of labels are used. Thus, this poses an excellent problem for unsupervised machine learning. Unsupervised machine learning is the field of machine learning which learns patterns from unlabelled data. A very commonly used unsupervised technique is called Clustering [51], which attempts to find the natural groupings in a dataset based on the similarity and dissimilarity between its data points. There exists no work in the current literature which presents an unsupervised approach to classifying rice grain damages.

1.5 Main Contributions

In this study, an unsupervised deep learning based method for a fine grained damage classification of rice is developed. In particular, this study addresses the two main concerns of: P1) Problems faced in manual labelling such as subjectiveness, erroneousness, overlapping damages, etc; P2) Task specific nature of a supervised machine learning algorithm which limits its usage to different scenarios.

The main contributions of this study are summarised as follows:

1) Development of a large dataset (20,134 images) of high magnification 24 Megapixel (6000 x 4000 pixels) images of individual rice grains (processed white rice) spread across different damages. In particular, BPT 5204 variety was used, which was collected from the rejected pile of a double colour Sortex machine (Milltec). The dataset very much stands as a representative for most varieties of processed white rice.

2) Use of a completely unsupervised (without any manual labelling) approach to classify the rice damages. This addresses point P1 and helps approach this problem without any preconceived notion of any particular groupings. In particular, contrastive self-supervised learning (subfield of deep unsupervised learning) is used to learn the representation of images. These feature representations learned are further reduced to two dimensions enabling much easier visual analysis. Finally, a clustering algorithm is implemented.

3) Success in clustering the rice grains into six main classes based on different damages, with establishing valid visual inspection parameters reasoning the clustering. Further, it is also discussed how the method can be extended to subclassify these damages according to the user's needs by providing a low-level control. This way it has been tried to address point P1 and P2 both.

To sum it up, in this study is presented a fine-grained damage classification method for quality evaluation of processed white rice based on deep unsupervised learning.

Chapter 2 - Related Work

2.1 Machine Vision for classification of rice grains

A typical machine vision system for food quality evaluation would follow five steps: Image Acquisition, Pre-processing, Segmentation, Feature Extraction, Classification [24]. Several past efforts have been made to classify rice using machine vision systems. These classifications have been based on a number of criteria. Grading of rice based on different size and shape [15] of the grains has been very common. For instance [25] classified long, short, slender, bold and round grades of rice based on variations in kernel size and shape. They applied a multi-class Support Vector Machine (SVM) on a dataset of 800 kernels. [26] classified rice grains into three grades; full, half and broken, using aspect ratio calculations. Over 1000 grains samples were used for each class. [27] used flatbed scanners to detect the number of broken kernels using image analysis.

A lot of the literature has been focussed on variety based classification. [28] classified different varieties in the Philippines; using variations in shape and size. [29] used colour and morphological features to classify rice into common rice, glutinous rice, rough rice, brown rice. In the literature, Supervised Machine Learning [21] based methods in particular Support Vector Machines [25, 31] have been commonly used for tasks with trivial feature extraction. [30] classified rice into three rice varieties (baldo, osmancik, yasemin) and broken using various machine learning models like Naive Bayes and Random Forest.

With the advent of deep learning [32], there has been a significant shift from classical machine learning to convolutional neural networks in a lot of computer vision applications [33-35]. [36] applied support vector machines (SVMs) and artificial neural networks (ANN) on a dataset of 600 images (90 test images) to categorise rice grains into three kinds; Ponni, Basmati, Brown. The ANN performed better with an accuracy of 93.34% compared to 92.2% of SVM. [37] demonstrated the efficacy of CNNs in feature extraction, which enabled the classification of rice grains of three different varieties (Japonica, Glutinous, and Indica) using a dataset of 3819 images (2854 calibration and 965 for validation) with an accuracy of 99.52%. Further, [38-40, 47] all have focused on variety classification and that using deep learning methods.

There have been fewer attempts in damage classification of processed rice grains. Those few include the work of [41] who classified rice grains into four classes (broken, chalky, red-spotted, and black-spot) using a dataset of 200 images (40 for algorithm development and 160 for its assessment). [31] classified the flawed red indica rice kernels (different from milled rice) into four classes, namely, cracked, chalky, damaged, and spotted, with an overall accuracy of 96.4%. The damaged rice kernels in the images were detected using a support vector machine (SVM) classifier. Another SVM performed the grey-level segmentation, which was subsequently used to extract the chalky areas. Damaged and spotted areas on the rice kernels were identified using edge detection and morphological methods such as dilation and morphological closing. The overall accuracy achieved was 96.4%. [101] used a deep CNN architecture EfficientNet to classify white rice into healthy, discoloured, broken, half-chalky, full-chalky, chalky-discoloured, and normal-damage. [42] developed an automated inspection system to classify brown rice into healthy, cracked, chalky, immature, dead, damaged, and broken classes, which were prepared artificially in a lab. [49] classified grains into four categories based on stress cracks. [50] tried to separate pecky kernels from healthy ones and further classified the pecky kernels into four categories based on size and location of damaged region. Some other similar attempts on damage classification other than rice grains include [43, 44] for wheat, [45] for triticale and [46] for corn. To our report, there has been no attempt at classification of rice damages in an unsupervised way.

2.2 Clustering of images

This study has been approached in a completely unsupervised way, with the initial goal to cluster the 20,134 images of rice into broader damage groups. Clustering [51] is an important data analysis tool, which attempts to find the natural groupings in a dataset (set of 'data points') based on their similarity and dissimilarity. Clustering as a concept is well studied in statistics and machine learning [66]. Datasets can be of numeric, textual or image-based.

It is found that clustering images is more challenging as compared to other forms of data, in particular due to the "curse of high dimensionality" [67]. Note that in the study, image clustering is referred to clustering with individual images as data points and is not to be confused with image segmentation [68] where individual pixels of an image are clustered. Most of the current research on image clustering is focussed on pixel wise segmentation and not

clustering with images as the data points. [69] provides a survey on image analysis through clustering. The basis of image clustering is how one measures the similarity among images (data points) and the discrepancy between semantic clusters.

Image Clustering typically follows two steps: 1) Feature Representation of images; 2) Implementation of a clustering algorithm [54, 65]. Since, directly clustering the high dimensional images is very challenging, the most important part becomes to design appropriate lower dimensional representations/features for the images. In this study, our aim was not only to cluster images into different damage but also be able to analyse various intra cluster patterns and possible sub clusters. For this, as shown in this study, it was necessary to be able to visualise the clusters and hence the need of two-dimensional embeddings of images. Thus, the approach as taken by us was a three-step approach: 1) Feature Representation of images; 2) Two dimensional embeddings via dimensionality reduction; 3) Visual Analysis of clusters.

Most of the present literature on feature representation of unlabelled image data focusses on manual feature extraction. For unlabelled data, features can't be extracted by training a Deep-CNN on a supervised task, as it requires labelled data. For feature extraction of rice grain images, the classical approaches [25, 31, 41, 42] use manual feature extraction based on factors like dimensions, contours, colours, shape and chalkiness. Methods like HOG [58], SIFT [59], are also used for general feature extraction. However, these features are not that robust and may lose representations from messy variables (e.g., background, rotation, brightness).

Recently, deep unsupervised representation learning has been widely explored in Computer Vision to learn informative feature representations of images. It aims to map samples/images into semantically meaningful representations without human annotations, which can facilitate various down-stream tasks, such as object detection, classification and clustering. These representations are found to be quite robust and can be made invariant to variables like brightness, orientation, etc. Recent advances in contrastive self-supervised learning [70] approaches for computer vision like SimCLR [60], SimSiam [61] and SwAV [76] have opened up new opportunities for learning visual representations without manual annotations. In natural image classification, these approaches provide comparable results to those obtained using supervised learning. Following its success in computer vision, this strategy has been adopted in several applications in other research fields including clustering of mass spectrometry imaging data [62], scRNA-seq data [63] and geo-located datasets [64].

In this study, a contrastive self-supervised learning approach based on [60] is used for learning the feature representation (128 dimensional) of images, followed by dimensionality reduction and clustering.

Chapter 3 - Materials and Methods

3.1 Data Acquisition

3.1.1 Sample Collection

The rice variety BPT 5204 was adopted in this study. It is a medium-sized grain that is mostly farmed in the states of Andhra Pradesh, Telangana, Karnataka, and parts of Madhya Pradesh, Bihar, and Uttar Pradesh. Rskissan Foods (India) Private Limited, Mirzapur, Uttar Pradesh, contributed to this research by allowing rice grain samples to be obtained from the rejection pile of a double colour Sortex machine (Milltec) for the 2019 growing season. A total of 16000 rice grains were collected, spread across various types of damages including 'healthy' grains. High magnification images of 24 MP (Megapixels) were taken for each grain. For some, both sides of the grains were covered, and a total of 20,134 coloured images of individual grain kernels were acquired as a data set.

3.1.2 Imaging System

The computer vision system used in the study consists of an LED light source for illumination because of its benefits such as long life, energy-efficient, and no heat or UV emissions. To produce high-resolution (6000 x 4000 pixels) and high-magnification (3.9 m/pixel) images of rice kernels, a mirrorless camera (Alpha5100, Sony) with a complementary metal-oxide-semiconductor (CMOS) sensor and paired with a microscopic zoom lens (Zoom6000, Navitar) was utilised (The resolution mentioned is in height x width format). A zoom lens (Navitar) which gives us the working distance of 10 cm with 4.5x magnification so that we can easily pick and place the rice grain and get the most amount of rice in the image, a computer system (i5-9500 CPU @ 3.00GHz, Dell OptiPlex 3070) and software (Sony, Imaging Edge Desktop).

As shown in Fig. 3, the camera is installed on a vertical platform. A 3-axis stage is positioned beneath the camera, on which rice grains are laid over a blue background, to capture the image.

The complete setup is contained in a rectangular box with a black inside coating to eliminate stray reflections, as well as an LED panel and high-efficiency diffuser on its sides, to provide a clean image. The camera is attached to the PC and is operated by software to shoot and save photographs without disturbing the 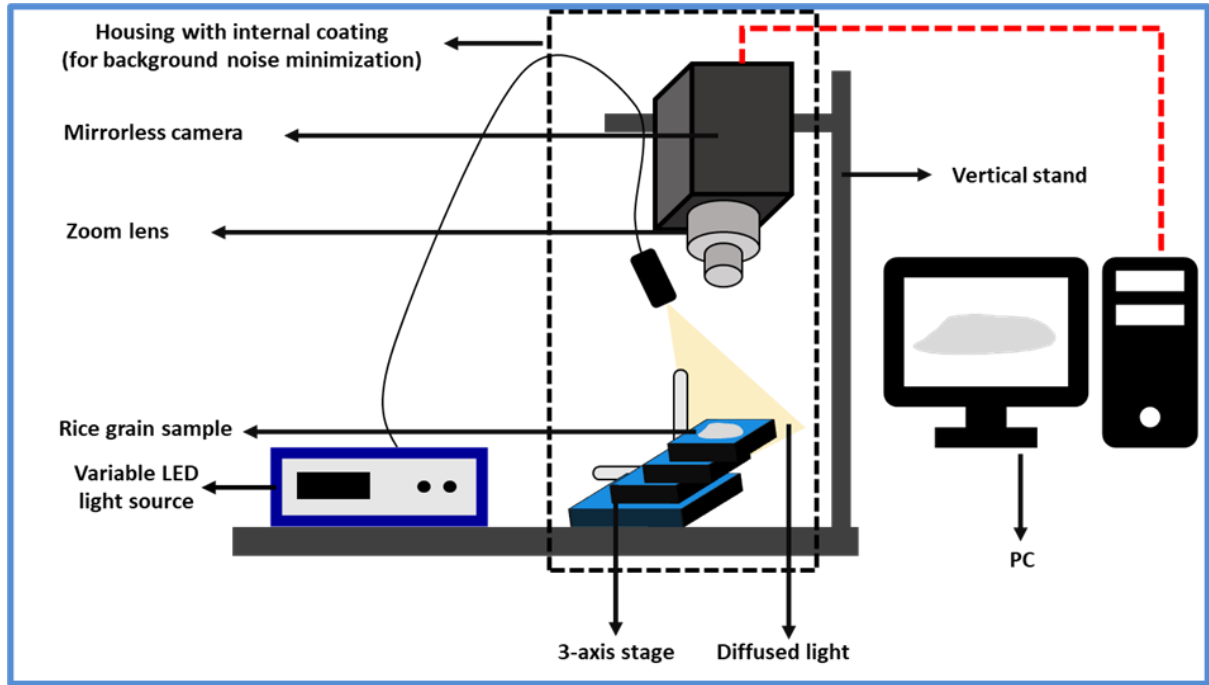setup, as shown by the red dotted line. The camera is attached to the PC and is operated by software to shoot and save photographs without disturbing the setup, as shown by the red dotted line.

Figure 3. Schematic representation of the image acquisition system

3.2 Dataset Details

The dataset contains a total of 20,134 images of individual rice kernels. All the images are in RGB format with a resolution of 6000 × 4000 (height × width) pixels i.e., 24 MP (Megapixels). In this study, for training the SimCLR model, the images were resized to 75 × 50 pixels. This was in order to be able to conduct feasible experiments, considering factors like the computer memory available, training time, etc. Although the final method uses 75 x 50 resolution images, the acquisition of high resolution (24MP) images was really helpful and much needed in the overall analysis of the clusters. It helped in understanding the distribution of various regions in the clusters by analysing the corresponding images. This proved to be

difficult when dealing with images of lower resolution (75 × 50), since in the study we focus on a finer classification of damages and the features are very complex.

3.3 The Contrastive-RC Method

The method used in this study is divided into three phases. To avoid confusion, it is collectively referred to as *Contrastive-RC* i.e., short for contrastive rice classification. Its three phases are as follows: 1) Representation Learning: A convolutional neural network (the 'encoder') learns a 128-dimensional vector representation for each image through a *contrastive self-supervised representation training* phase; 2) Dimensionality Reduction: A dimensionality reduction technique UMAP is used to reduce the dimensions of these embeddings from 128 to 2, which are then plotted on a 2D plane. Each point in the 2D plot represents a distinct rice grain (image); 3) Cluster Analysis: The final two-dimensional embeddings are then clustered using HDBSCAN and analysed for further patterns/subclusters. Fig. 4 shows a schematic representation of the method after training.
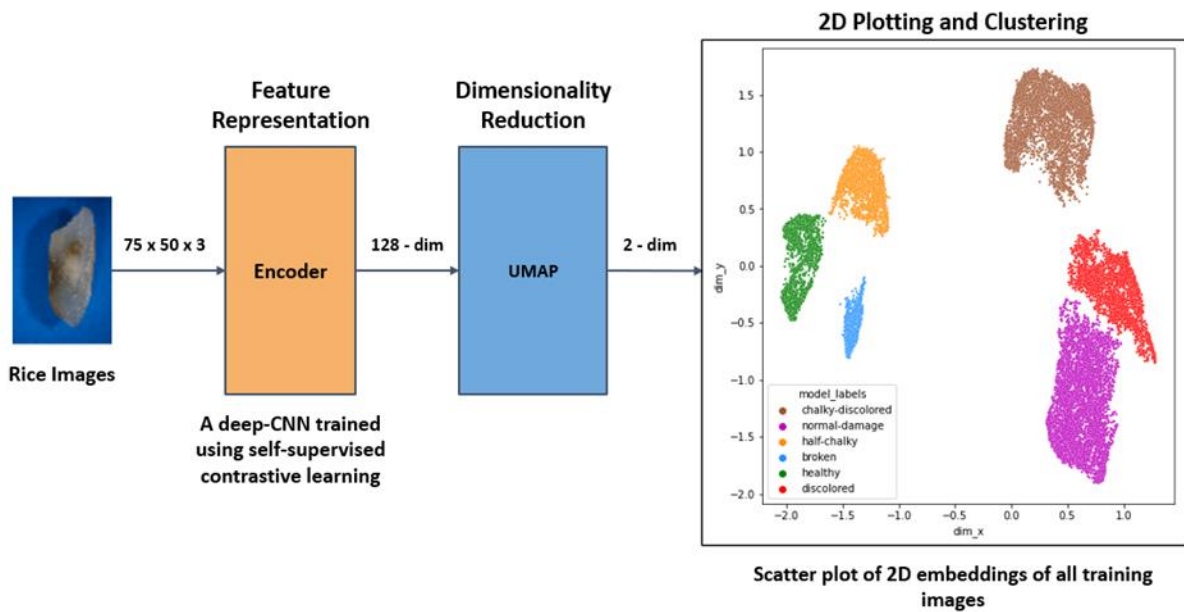

Figure 4. A schematic representation of the method Contrastive-RC after training: 1) All the images are passed through a trained convolutional neural network (the 'encoder') to represent the images as 128 dimensional vectors. The encoder is trained using a contrastive self-supervised representation learning phase; 2) A dimensionality reduction technique UMAP is used to reduce the dimensions of these embeddings from 128 to 2; 3) The 2D embeddings are then plotted and clustered using HDBSCAN.

3.3.1 Feature Representation Learning using SimCLR

Contrastive self-supervised learning [70] is a technique used for representations learning, that aims to map data into semantically meaningful representations; The goal of contrastive representation learning is to learn such an embedding space in which semantically similar sample pairs of data points stay close to each other while dissimilar ones are far apart. The technique has found widespread application in the field of computer vision for learning representations (embeddings) for unlabelled images. It is an unsupervised method which uses self-supervision and creates pseudo labels for itself, to learn the representations which can then be used for other downstream tasks like classification, clustering, segmentation, etc. These representations can facilitate various down-stream tasks, such as object detection, classification and clustering. Contrastive self-supervised learning has produced state-of-the-art results in semi-supervised and unsupervised settings [71, 72].

On a high level, it tries to learn lower dimensional representation of images that are invariant to image augmentations [73]. It does this by pulling augmented versions of the same image (positive pair) closer while pushing the augmented versions of different images (negative pair) farther from each other. The whole idea of augmentation is to create two versions of the same image, such that the versions are different in appearance but have not lost their true semantic meaning. This is what the model tries to learn. In this study, the architecture used is based on *SimCLR* [60] which stands for "A Simple Framework for Contrastive Learning of Visual Representations". It has produced state-of-the-art results on many tasks. The method (Fig. 5), as novel and powerful it is, is fairly trivial to understand.
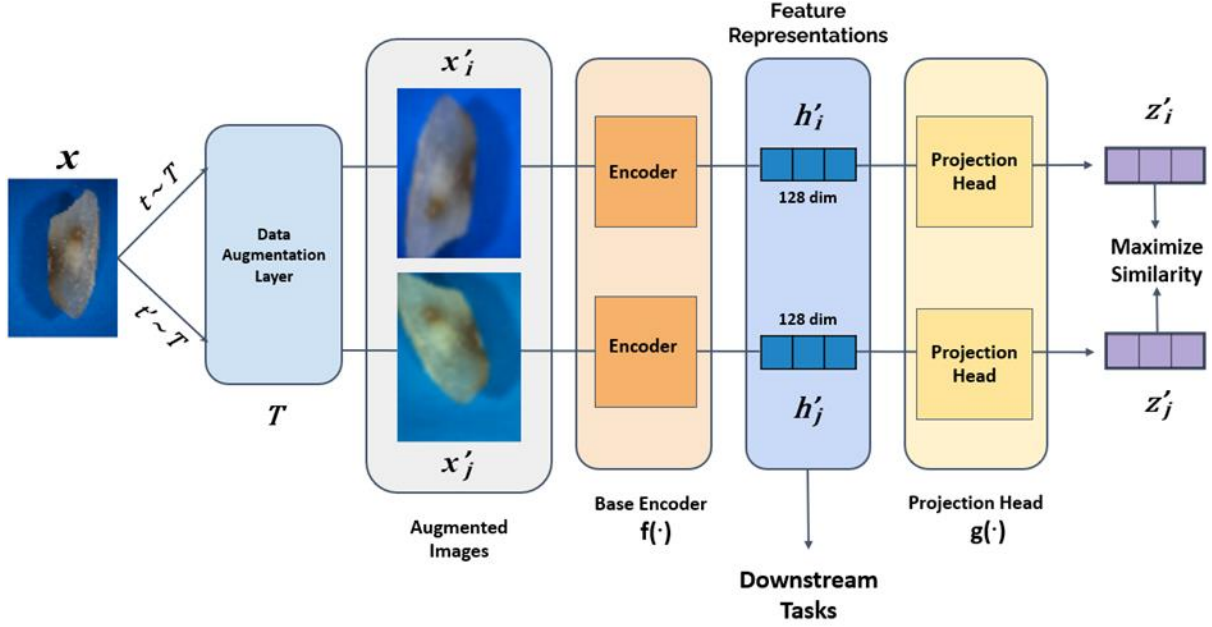
Figure 5. Schematic representation of the SimCLR method used for learning visual representations of the rice images in phase 1. The feature representations h_i and h_j are the ones finally used for the clustering task.

The steps involved in the algorithm implementation are explained below. Afterwards, the same are summarized mathematically in Algorithm 1.

1. Firstly, batches of randomly selected images are drawn from the image dataset D. These batches are fed into the model one by one. For each batch, the first step is to create a pair of different augmented versions x'_i and x'_j of each image x in it, through an augmentation function T. Here $x'_i = t(x)$ and $x'_j = t'(x)$ where $t', t \sim T$.

2. All the augmented image pairs are passed through a deep convolutional neural network ('the encoder') $f(\cdot)$ which encodes them into lower dimensional vector representations. Here, $h_i = f(x'_i)$ and $h_j = f(x'_j)$ where $f(\cdot)$ represents the global averaged output of the last convolutional layer before the fully connected layers start. These representations h_i and h_j are the ones which are used after training for the downstream tasks (clustering in our case).

3. The encoded representations h_i and h_j of the image pairs are then passed through a nonlinear projection head $g(\cdot)$ which is a simple neural network of two dense layers (128 dimensions each) with a relu activation in between. It produces the vectors z_i and z_j for each pair of images. Here, $z_i = g(h_i) = W^{(2)}\sigma(W^{(1)}h_i)$ where σ represents the non-linear relu function and $W^{(1)}, W^{(2)}$ represent the weight matrix of the first and second dense layer respectively. The vectors h and z are 128 dimensional long. The projection head is added as it proves to increase the representation quality of the layer before it (h_i and h_j). The projected embeddings are finally used by the model to optimise its parameters. It tries to bring closer the embeddings for the

14

augmented versions of the same image and push farther apart the embeddings coming from augmented versions of different images.

4. This is achieved by minimising a loss function called "NT-Xent loss" (Normalised Temperature-Scaled Cross-Entropy Loss) which is based on cosine similarity between the embedding pairs. Say a batch of N images are fed to the model to produce $2N$ projected vectors (N pairs of z_i, z_j). Let $sim(z_i, z_j) = z_i^T z_j / \|z_i\| \|z_j\|$ represent the cosine similarity between z_i and z_j, then the loss function for a positive pair of examples (i, j) is defined by Eq.3, where т is an adjustable temperature parameter allowing to scale the inputs and widen the cosine similarity range. The final loss function Eq.2 is the average of losses over all the points. This loss function is then minimised by updating the parameters of f and g using gradient descent with Adam optimiser [75].

5. Finally, the model f is retrieved for feature representation, and g is discarded.

Algorithm 1: SimCLR's main learning algorithm

$\boldsymbol{Input} : D, T, N, т, f, g$
$\boldsymbol{for}\ batch\ B\ \in D\ s.t\ B = \{x_k\}_{k=1}^N$
    $\boldsymbol{for}\ k\ \in \{1, \ldots \ldots N\}$
        $x'_{2k-1} = t(x_k)\ \&\ x'_{2k} = t'(x_k)$
        $h_{2k-1} = f(x'_{2k-1})\ \&\ h_{2k} = f(x'_{2k})$
        $z_{2k-1} = g(h_{2k-1})\ \&\ z_{2k} = g(h_{2k})$
    $\boldsymbol{for}\ i \in \{1, \ldots \ldots 2N\}\ \boldsymbol{and}\ j \in \{1, \ldots \ldots 2N\}$
        $sim(i,j) = z_i^T z_j / \|z_i\|\|z_j\|$ *(1)*
    $\boldsymbol{Minimize}\ L = \frac{1}{2N} \sum_{k=1}^{N} [l(2k-1, 2k) + l(2k, 2k-1)]$ *(2)*
        $by\ updating\ f\ \&\ g$
    $\boldsymbol{where}\ l(i,j) = -log \dfrac{exp(sim(i,j)/т)}{\sum_{k=1}^{2N} 1_{|k \neq i|}\ exp(sim(i,k)/т)}$ *(3)*
$\boldsymbol{return}\ f$

There are a bunch of hyperparameters that can be tuned to improve the model's performance. These include the batch size, number of epochs and the temperature т. Further, the augmentation layer T, and the architecture of the encoder $f(\cdot)$ are also variables which needs to be set accordingly. Here, batch size becomes a more important parameter than usual since (loosely speaking) the objective can be interpreted as a classification over a batch of images. In the paper [60], larger batch sizes (8192) and longer training (1000 epochs) have produced better results. Scaling up of the encoder also has proven to increase the performance. Hence, in our study there was a trade-off between using a larger architecture and a larger batch size dude to memory limitations. After a lot of experimentations, MobileNetV3 [74] was chosen as the encoder with a batch size of 2048. A detailed discussion on the hyperparameter tuning is provided in section 4.2.

3.3.2 Augmentation Layer (T) in SimCLR

The augmentation layer is a very important part of the whole model since it provides the basis of what the model tries to learn. The augmented versions of an image are supposed to be created in a way such that although different in appearance, the true semantic meaning of the images remain the same i.e., in essence one is able to tell that they come from a similar distribution. For example, changing the shape, size and orientation of a discoloured grain will not change the fact that it is discoloured. It is shown that stronger data augmentations boost the performance significantly [60, 61, 76].

The augmentation pipeline used in the model is shown in Fig. 6. As a pre-processing step, the images first undergo rescaling of pixels by a factor of 1/255 in order to normalise the data. Then random flipping is used in both horizontal and vertical directions. Random rotation is performed in either direction with a maximum angle of rotation being 18 degrees.

The most important part of the augmentation pipeline was cropping and brightness alteration. As discussed before, the augmentations had to be performed carefully such that the real meaning of the images don't change. For instance, a large increase in brightness can make a healthy grain seem like its chalky, or too small of an image crop may lose the part of the grain that is damaged hence defeating the whole purpose. Keeping these heuristics in mind, the types of augmentation were chosen. Both cropping and brightness alteration were performed randomly. For cropping, a parameter *min_area* was used which signified the minimum fraction of area that needs to preserved in the image after cropping. Cropping was performed with the help of random translation and random zoom. This also brought a relative change in height and width of the grain which ensured that our model could be robust to aspect ratio changes. For adjusting brightness, a brightness factor *alpha* was used which represented the max amount by which the brightness was increased (or decreased) i.e., the maximum value which was added (or subtracted) to the image pixels (after rescaling). Both the parameters min_area and alpha underwent considerable tuning (as discussed in section 4.2) and were finally set to 0.6 and 0.4 respectively.
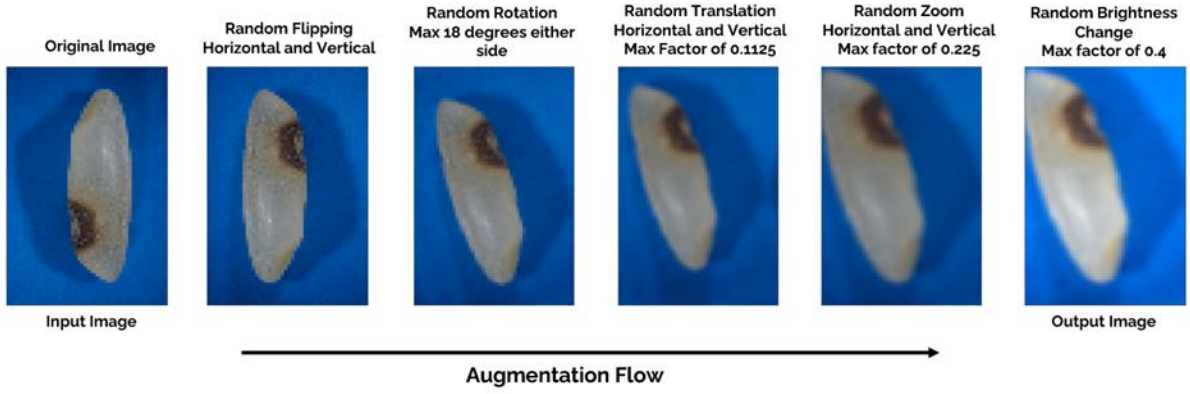
Figure 6. Augmentation Pipeline

Feature representations can also be learned in a supervised setting (labels available), however that is through training on hard coded labels. For example, the CNN model would be instructed to encode 'mildly' discoloured grains same as 'strongly' discoloured grains, or to encode 'slightly' broken grains same as 'largely' broken grains. Thus, although the method may work well to separate the damages, it may not be able to capture the real essence of the images as a whole. SimCLR, on the other hand, learns representation in a highly intuitive way that is not limited by hard coded labels and has the ability to capture diverse patterns. By training on an infinite number of pseudo labels, it captures the true semantic meaning of all the image's attributes. As a result, it is not constrained by a specific scenario and may capture many more significant representations than are invariant to such data augmentations.

3.3.3 Dimensionality Reduction and Clustering

The representations learnt from the SimCLR model are 128 dimensional long. These can be clustered directly; however, it is not possible to visualise 128 dimensional points and hence their analysis becomes difficult. Thus, the dimensions were reduced from 128 to 2 (phase 2). This way, the embeddings could be plotted in 2D space and analyse the results visually in a much more intuitive way. For this, a dimensionality reduction technique UMAP [77] was used. It is much faster as compared to other techniques like PCA (Principal Component Analysis) and TSNE [87] still producing high quality results and was very suitable for experimentation purposes. For instance, it is ~ 45 times faster than TSNE on the 784-dim MNIST dataset [100]. UMAP tries to learn the manifold structure of the data and create a low dimensional embedding that preserves the essential topological structure of that manifold.

17

The important parameters include *min_dist* and *n_neighbors*. The parameter min_dist signifies how closely the data points are packed in the lower dimension. It is the minimum distance apart the points are allowed to be in a lower dimensional space. Further n_neighbors signifies the size of the local neighbourhood to consider for each point when attempting to map the points to lower dimensions, and is an important parameter in balancing the local vs global structure. It is advised to try multiple values for both these parameters and choose the ones around which stable embeddings are realised. Finally, the 2D embeddings are clustered using HDBSCAN [65] (phase 3), a hierarchical density-based clustering algorithm. The important parameters include *min_cluster_size* and *min_samples*. The parameter min_cluster_size refers to the least number of points needed for a grouping to be called a cluster, whereas min_samples refers to the minimum number neighbours to a core point.

Chapter 4 - Experiments

4.1 Experimental Setup

All experiments were carried out in the Google cloud environment using a 64-bit Debian GNU/Linux operating system powered by an Intel (R) Xeon (R) CPU @ 2.20 GHz and 26 GB RAM, with NVIDIA Tesla P100 containing 16 GB memory. All the code is implemented in Python. *Keras* [80] with *Tensorflow* [81] backend was used for the implementation of SimCLR model. Official python libraries of UMAP [82] and HDBCAN [83] were used for their respective implementations.

4.2 Hyperparameter Tuning and Selection Criterion

The study, as previously stated, was conducted in three phases. Each phase has its own set of hyperparameters that need to be tuned to get the desired results. The first two phases were the most important in terms of tuning; phase 3 required minimal tuning. Phase 1, or representation learning phase, consisted of hyperparameters such as the *num_epochs* (number of epochs), batch size, temperature, and (*min_area*, *alpha*) from the augmentation layer. Phase 2, in which the 128-dimensional representation vectors were embedded in 2D space, included the hyperparameters *min_dist* and *n_neighbors*. The hyperparameters for phase 3 (Clustering)

comprised *min_cluster_size* and *min_samples*. Hence, the hyperparameters were separated into three sets.

Tuning in both phase 1 and 2 was performed based on visual analysis of the 2D embeddings. Although phase 1 was independent of phase 2, it was tuned by feeding its outputs (128-D feature representations) through the second phase and examining the clusters/groupings in the 2D embeddings. The main idea was to select the hyperparameters which could give the best clustering; well separated groups of points in the 2D space.

Metrics like *silhouette score* and *calinski-harabasz index* [97] were not used for clustering evaluation as these metrics depend on the performance of the clustering algorithm and the labels provided by it. And it was found to be quite difficult to fit the clustering algorithm on every instance so as to provide the optimal labelling. Since the clustering were observed in two dimensions, visual analysis proved to be a highly effective approach to evaluate the results. The basis of selection was dependent on how well separated the clusters were. Further, the points corresponding to each cluster were analysed to provide proper justification for the separation. Possible subclusters in a cluster gave the indication of possibility of further separation, which also was a considerable factor in determining further experimentations and selections.

Many experiments were carried out using various combinations of all the hyperparameters. The most important ones were the num_epochs and the augmentation parameters (min_area, alpha). We first tuned the other hyperparameters to keep them constant for further experiments. The second phase hyperparameters min_dist and n_neighbors were the easiest to configure. The optimal values were chosen as those around which stable embeddings were realised. The values of 0.001 and 50 respectively were found to be optimal. These values were found to be independent of the hyperparameters used in the first phase and were kept constant for all the subsequent experimentations.

In the first phase, batch size was one of the important parameters and suggested to be kept at larger values [60]. Through experimentations, a batch size of 2048 was found to be optimal. One disadvantage of using larger batch size was that it restricted us in the encoder model size, which also shown to perform better when larger in size. Hence, MobileNetV3 was chosen which is one of the smallest CNNs in the literature and still compares in performance with the larger state-of-the-art models. The optimal temperature was found to be. For the augmentation

layer, min_area and alpha were tuned. The set of values tried were {0.5, 0.6, 0.7, 0.8, 0.9, 1.0} and {0.0, 0.1, 0.2, 0.3, 0.5 0.4} respectively. The most optimal values for both were found to be 0.6 and 0.4 respectively.

The parameter num_epochs was not that straightforward to set and a wide range had to be tried for each set of hyperparameters. In general, the models were trained for 300 epochs, or until they achieved a contrastive accuracy greater than 99.5 percent, while also checkpointing (storing) the intermediate versions. Contrastive accuracy is a self-supervised metric that represents the proportion of cases in which the image's representation is more similar to that of its differently augmented version than to that of any other image in the current batch. This was only used as a guideline for when to end training and did not provide a way to choose the ideal epoch. For each epoch, the saved models were used and the corresponding clusters were observed to find the best epoch. Typically, it was found to be between 100 and 200. The most optimal clustering achieved is shown in Fig. 7. The corresponding hyperparameter values were 184 for num_epochs, 2048 for batch size, 0.1 for temperature, (0.6, 0.4) for (min_area, alpha), 0.001 for min_dist, 77 for n_neighbors, 3 for min_samples and 500 for min_cluster_size.

For the experiments, out of the total 20,134 images, 19330 were kept for training and the rest 804 as test sets. The purpose of the test set was to ensure that the model did not overfit to the training data and could generalise to previously unseen data. Since there was no proper metric used, the evaluation on test was also done visually; For each test embedding, it was checked whether it belonged to the right cluster and whether the images corresponding to its neighbourhood points were similar to it or not. The test set was also utilised to evaluate the model's performance under different scenarios like grain rotation, translation, brightness change, etc.

Chapter 5 - Results and Discussion

The study was based on a large number of experimentations and simultaneous analysis of the 2D embeddings. The most optimal clustering was selected to be as shown in Fig. 7. In section 4.2, the methodology for selecting the right clustering was discussed. In this section, a detailed analysis of the final clustering is presented. Section 5.1 presents a broader analysis on properties of each cluster and how they are supported by the literature. On the basis of the clustering achieved, a total six damage-based classes (including the *healthy* class) are

presented. Section 5.2 presents an analysis of the method Contrastive-RC. We discuss its robustness towards changes in messy variables like brightness, grain orientation, etc. and its versatility for various use cases. Finally, in section 5.3, we present a more in-depth analysis of each cluster/class by examining different regions in each cluster. We also discuss the possible extension of the method for further subclassification.
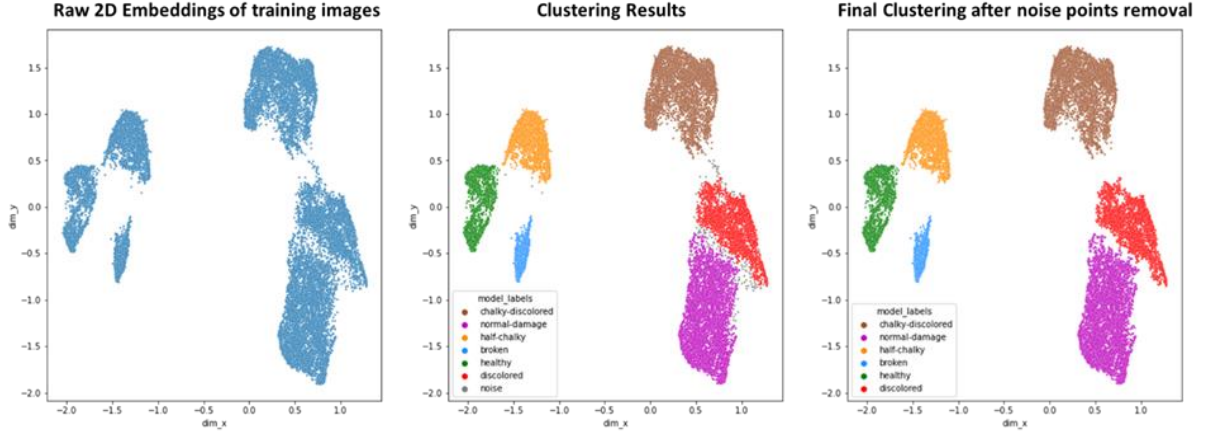

Figure 7. Left: Raw 2D embeddings of all the training images (total 19330). Middle: Results after clustering with HDBSCAN. Right: Clustered 2D embeddings after removing noise points (total 19179)

The final clustering can be seen in Fig. 7. It should be noted that each point (x, y) in the scatter plot represents a unique rice grain (image). As stated previously, out of the total 20,134 images, 19330 were kept for training and the rest 804 as test sets. The leftmost plot in Fig. 7 depicts the raw 2D embeddings of 19330 (training) images, which were obtained as outputs of phase 2. These were clustered using HDBSCAN and the results are shown in the middle plot. A total of six clusters are observed. HDBSCAN is a density-based algorithm that, in addition to assigning points to each cluster, also assigns some points to be "noise" points i.e., the isolated points that do not belong to any cluster. These are marked with grey colour in the middle plot. There were only 151 points which were marked as noise. For easier analysis, these were removed and the rest 19179 points are shown in the right most plot.

5.1 Inter Cluster Analysis

For each cluster the corresponding points/images were analysed. Here, the need for high magnification images was realised as the details in the images were very intricate, especially

for understanding different regions within each cluster. Each cluster was found to have a well-defined set of properties and was assigned a separate class name. Fig. 8 depicts the final clustering along with the class names and identifier for each cluster. The identifiers are the abbreviations of the class names assigned to them. Note that for the rest of the study, the grains of each cluster are referred to by the corresponding class name. The figure also shows a few image examples from each cluster. The number of points/images for each class in descending order are: 5599 for *normal-damage*, 4987 for *chalky-discoloured*, 3215 for *discoloured*, 2386 for *half-chalky*, 2061 for *healthy*, and 931 for *broken*.
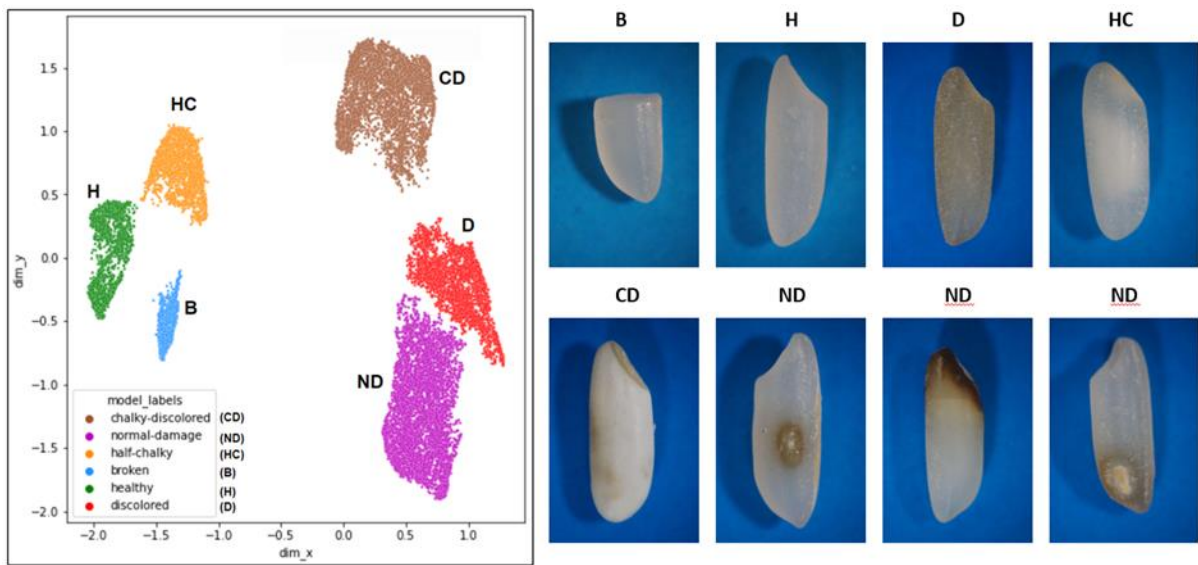

Figure 8. Final Clustering along with the class names and identifiers assigned to each cluster. A total of six clusters are observed. The identifiers are abbreviations of class names: 'H' for healthy, 'HC' for half-chalky, 'B' for broken, 'CD' for chalky-discoloured, 'D' for discoloured, and 'ND' for normal-damage. Along side are attached a few examples from each cluster. The three examples shown for normal damage class correspond to watermark, heat, and pin damage respectively from left to right.

Cluster H: "Healthy" class

The images corresponding to cluster H signify "healthy" grains. These grains are characterised by their clear translucent nature and have no type of surface damage. Further, most of the grains (>95%) were found to be full-sized. These are considered to be the highest grade of grains as translucent nature of the endosperm is a highly desirable characteristic and consumers pay a premium price for it [88].

Cluster B: "Broken"

The cluster B contains broken grains. Broken grains are generated as a result of internal stresses that are generated during or after the milling process, which causes the healthy grains to break down into tiny fragments. The grains in the cluster differ in sizes and can also come along with various damages. Literature shows that identification of broken rice and its size are an important factor in rice grading. Head rice recovery (HRR) is a commonly used attribute associated with broken kernels. It is defined as the proportion of paddy rice that retains 75% of its length after milling. It is a key rice grading attribute that has a significant impact on the rice market prices: higher the HRR; higher is the market value of milled rice [85, 86]. The broken grains can further be separated into second heads (larger sized broken), screenings (smaller sized broken) and brewers (very small sized broken) according to the market needs [87, 95]. Typically speaking, the broken grains without any other damage are the only ones used for human consumption. Damaged broken grains are considered unfit for human consumption and used primarily either as poultry feed or as a raw material for bio-ethanol production, etc.

Cluster HC: "Half-Chalky"

Chalkiness [98] is a well-known attribute in rice grains. Chalky grains contain opaque white or chalk-like regions. The degree of chalkiness for these grains is defined as the fraction of area that is covered by the chalky region. In cluster HC, chalky rice grains were observed but with lower degree of chalkiness; either the chalkiness was mild or if prominent (completely opaque) then it only covered a part of the grain and not the whole. In most of the grains, about half of the total surface area was observed to be chalky and hence they are termed as half-chalky.

Cluster CD: "Chalky Discoloured"

The grains of cluster CD are characterised by two prominent features namely chalkiness and discoloration (yellowish or brownish). Majority of the grains possessed chalkiness along with discoloration. The ones without (or very mild) discoloration had the full region covered in chalkiness. A few grains containing unremoved layers of husk were also observed in this cluster.

Literature shows that chalkiness is an important factor in rice grading. It is mostly associated with low quality of rice, which has a negative effect on the market price [88, 89, 98]. Chalkiness is negatively correlated to HRR i.e., HRR decreases with increased chalkiness [90]. Higher the degree of chalkiness, lower the market acceptability [88]. Thus, half-chalky grains would more readily be used for consumption in higher grades of rice than the chalky-discoloured ones. Attempts have been made to define and subclassify chalky grains but there is no consensus. Some classify grains as chalky only when more than 50 % of it is chalky [91]. The IRRI defined a scoring system from scale 0-5, 0 standing for no chalkiness and 5 standing for chalkiness greater than 80% [89].

Cluster D: "Discoloured"

The grains in cluster D were observed to only contain discolorations (brownish or yellowish), and were not identified with any kind of chalkiness or other damages. The degree of discoloration varied across grains. Some grains were observed to have very mild layers of husk along with the discolouration. Grain discoloration can occur in unsafe storage conditions and have a major impact on the market value [5].

Cluster ND: "Normal Damage" (Pin, heat, watermark)

Cluster ND grains were observed to contain three types of damages: *pin*, *heat* and *watermark*. Most of the grains were characterised by what we term as pin damage. These are characterised by ovular shaped indentations on the surface mainly caused by insects. The indentations range in size and colour (black, brown, white). Watermark damages are present in the form of water stains/spots on the grains. The third kind of damage found was *heat damage*. These grains are characterised by dark black regions seemingly burnt due to excess heat during processing. These kinds of grains are unfit for human consumption and typically used for other purposes like bio-ethanol production, poultry feed etc.

5.2 Analysis of the method Contrastive-RC

This section presents an analysis of the method used and how it is able to capture semantic representations of the rice images. As discussed previously, SimCLR is not bounded by hard coded labels and tries to capture a more general and meaningful representation of the images as compared to the ones in a supervised setting; SimCLR embeds semantically similar images near each other in the embedding space. Further, UMAP is able to reduce the dimensionality with also preserving the local and the global structure. On analysis of the predictions on test images, no overfitting was observed and the predictions were observed to be accurate; every test image was embedded in the space surrounded by semantically similar training images. Fig. 9 shows the distribution of the test embeddings as predicted by the model. Fig. 10 shows three examples of test images (normal, half-chalky, and discoloured) and the corresponding five nearest training images in the embedding space. The visual similarity is evident between them. Further, the specific augmentations used in the model have made it particularly robust to changes in brightness and grain geometry. We altered the brightness for images in the validation set and noticed how the embeddings of the original and altered version varied. It was observed that in most cases, the embeddings of both were very similar to each other meaning that the embeddings were very close in the 2D space. In the cases where embeddings were not similar, they were still observed to be in the same cluster. In particular, for alteration in brightness by factors of 0.1, 0.2, and 0.3, the percentage of the total altered versions (804) that were predicted to be in the same cluster as the original image were 95.77%, 88.43% and 80.22%. Furthermore, the method was robust to change in grain geometry: vertical and horizontal orientation, rotation, size, aspect ratio and position of the grain on the image. Fig. 11 shows three examples of test images and corresponding five randomly augmented versions. Below each image, is their embedding as produced by the model in the 2D space. Notice how the embeddings of all the augmented versions are very close the original image

Thus, it demonstrates the robustness of the method towards messy variables like brightness, orientation, etc., which are very commonly observed in real world scenarios. Further, notice that in the augmentation layer, the way cropping is performed, it also changes the aspect ratio of the images and in turn the rice grain. This makes the model robust to change in grain size and its aspect ratio i.e., the model will not be affected by changes in grain size and aspect ratio: longer, shorter, fatter and thinner version of the same type of grain will be embedded very similarly to each other. This allows the model to be easily extended to other varieties of

processed white rice. As far as the inference time of the model is concerned, the model is considerably fast, particularly when images are fed in batches. For batches of size 1, 64, 256, 1024, 2048, 4096 and 8192, the average prediction times were 2.11 s, 2.25 s, 2.40 s, 3.57 s, 5.53 s, 11.67 s, and 26.20 s respectively. The corresponding frequencies are 0.47, 28, 106, 286, 370, 351, 312 images per second. Here 's' is the unit in seconds and the time is the average taken over 1000 iterations. This includes the time from feeding the images to the model to the final class predictions. Hence, the model is able to process 370 images per second for a batch size of 2048, hence making it very suitable for real world application where fast processing is needed and images.
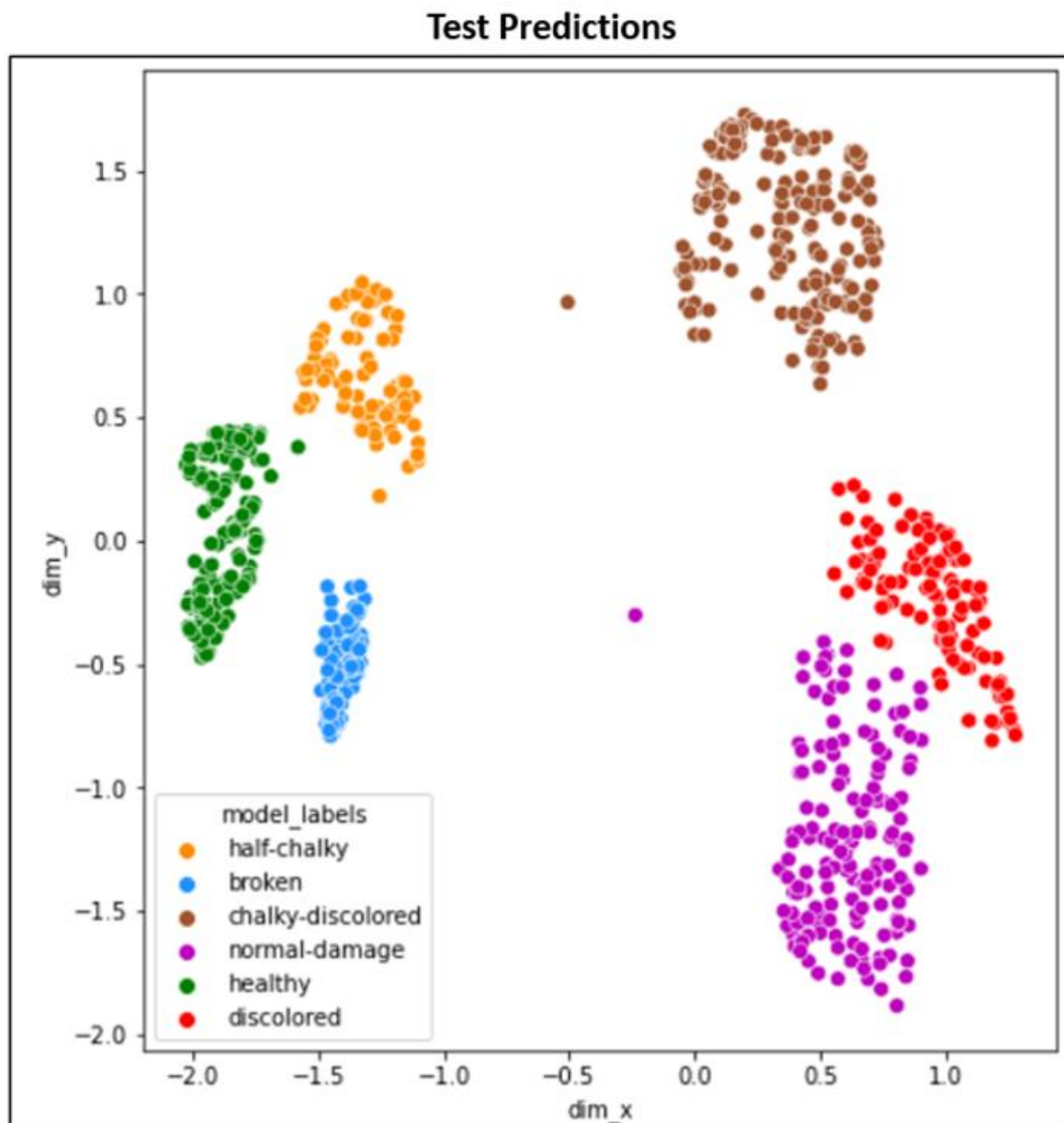

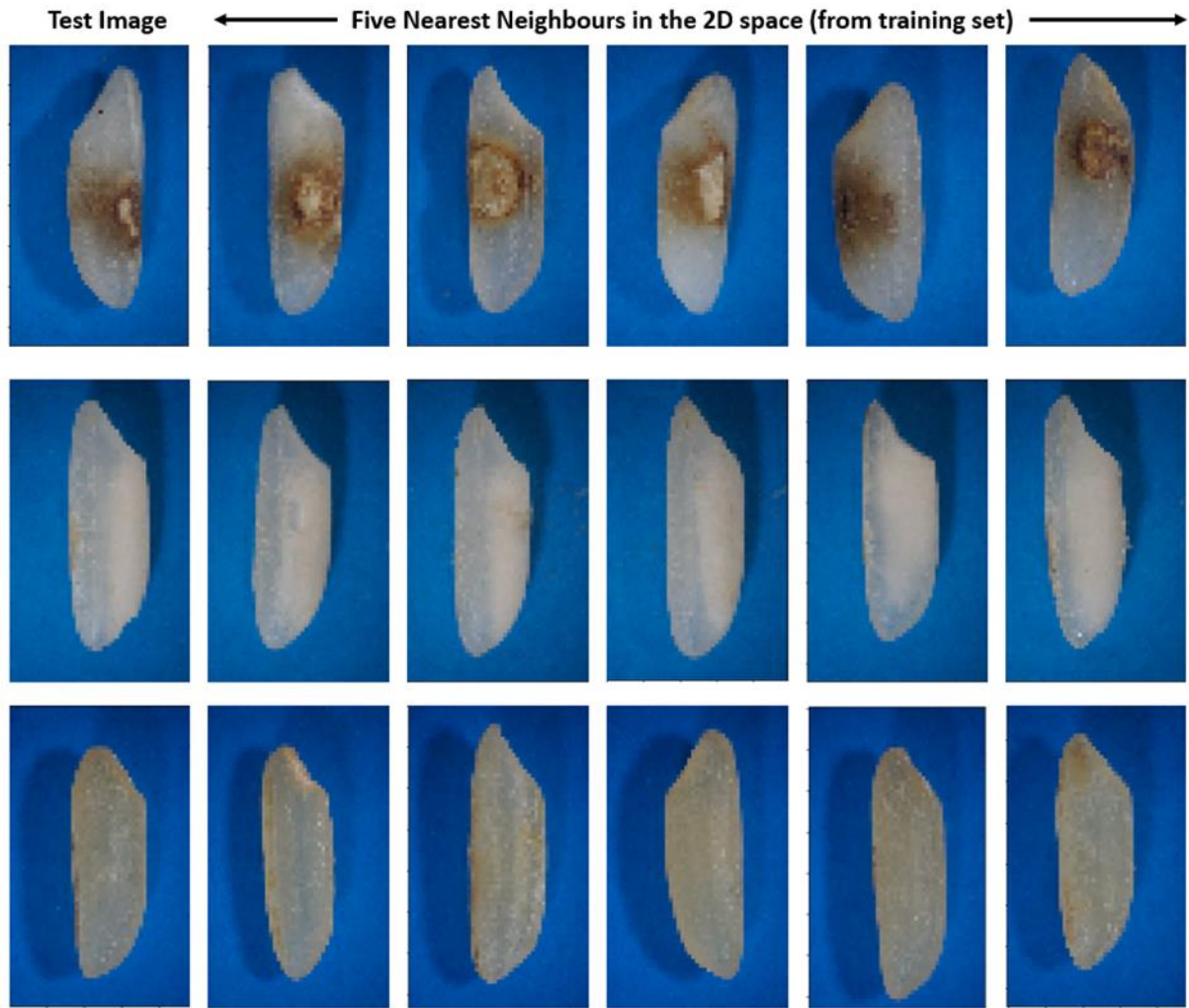*Figure 9. Embeddings of the test images (804) as predicted by the model*

0

Figure 10. Test images and corresponding five nearest neighbours in the embedding space. The five nearest neighbours to a test point signify the five closest points to it. The 'closeness' is measured by the Euclidian distance between them. Here the nearest neighbours are taken from the training set.
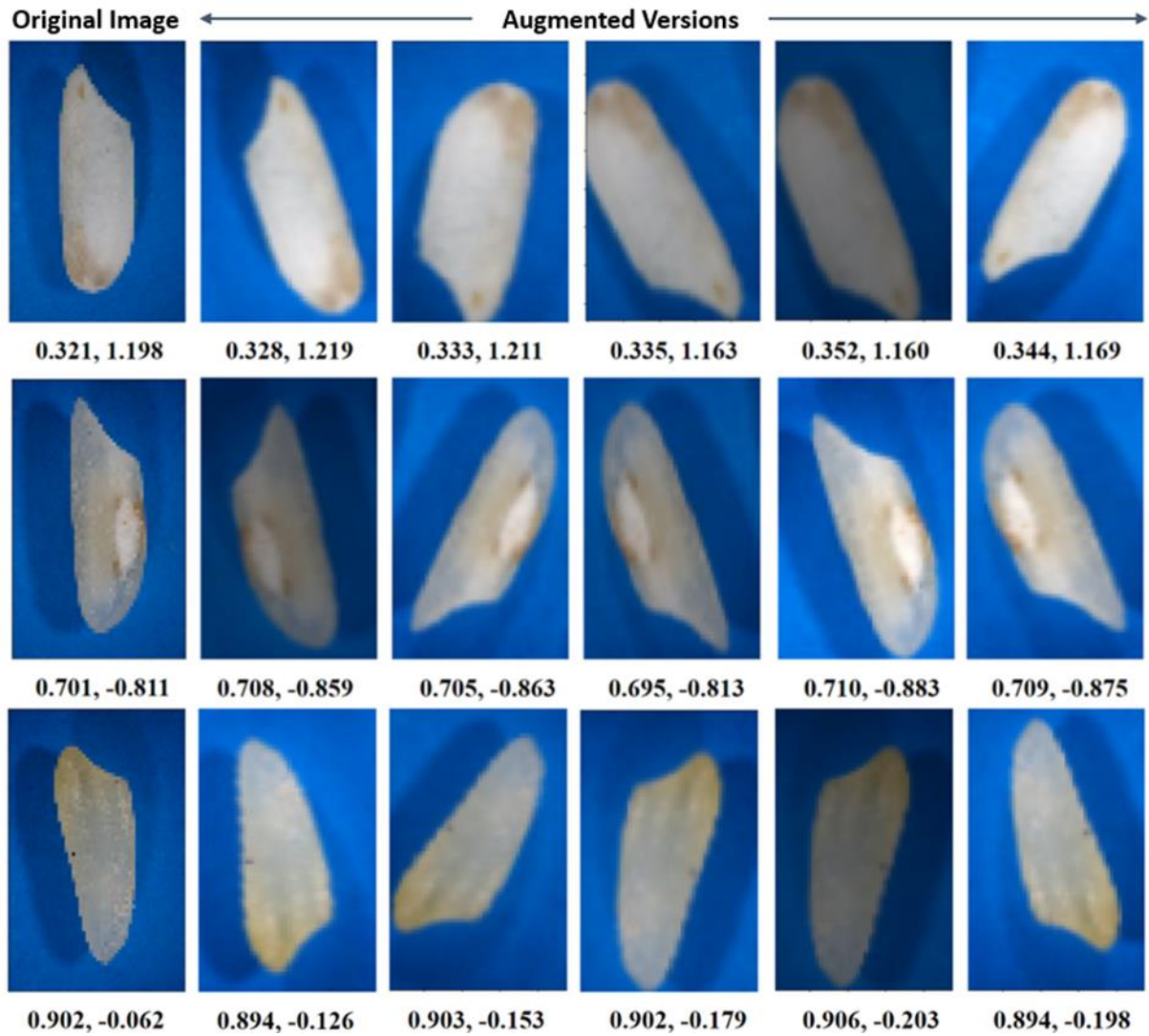
Figure 11. Test images (left most) and their corresponding augmented versions. Below each image is its 2D embedding as predicted by the model.

5.3 Intra Cluster Analysis

The ability of the SimCLR model to embed semantically similar images closer to each other in the 2D spaces allows for possible sub clustering. The images in different regions of each cluster were observed and very meaningful and information rich patterns in the distributions were found. Mostly, it was observed that there were no clear demarcations/boundaries as the distribution for each cluster was continuous in nature. For instance, discoloured grains did not have a further demarcation between say strongly discoloured and mildly discoloured because

there was no proper boundary to distinguish one from another; there was a continuous movement of grains from mild to strong discolouration. Fig. 12 shows different regions for each cluster and the corresponding grains. Each image is marked with the type of damage and the region it belongs to. For instance, 'CD4' corresponds to a *chalky-discoloured* grain belonging to the 4th region as marked in the cluster for chalky-discoloured. Further from now on, for each damage, region 1 is referred to as r1, region 2 as r2 and so on.
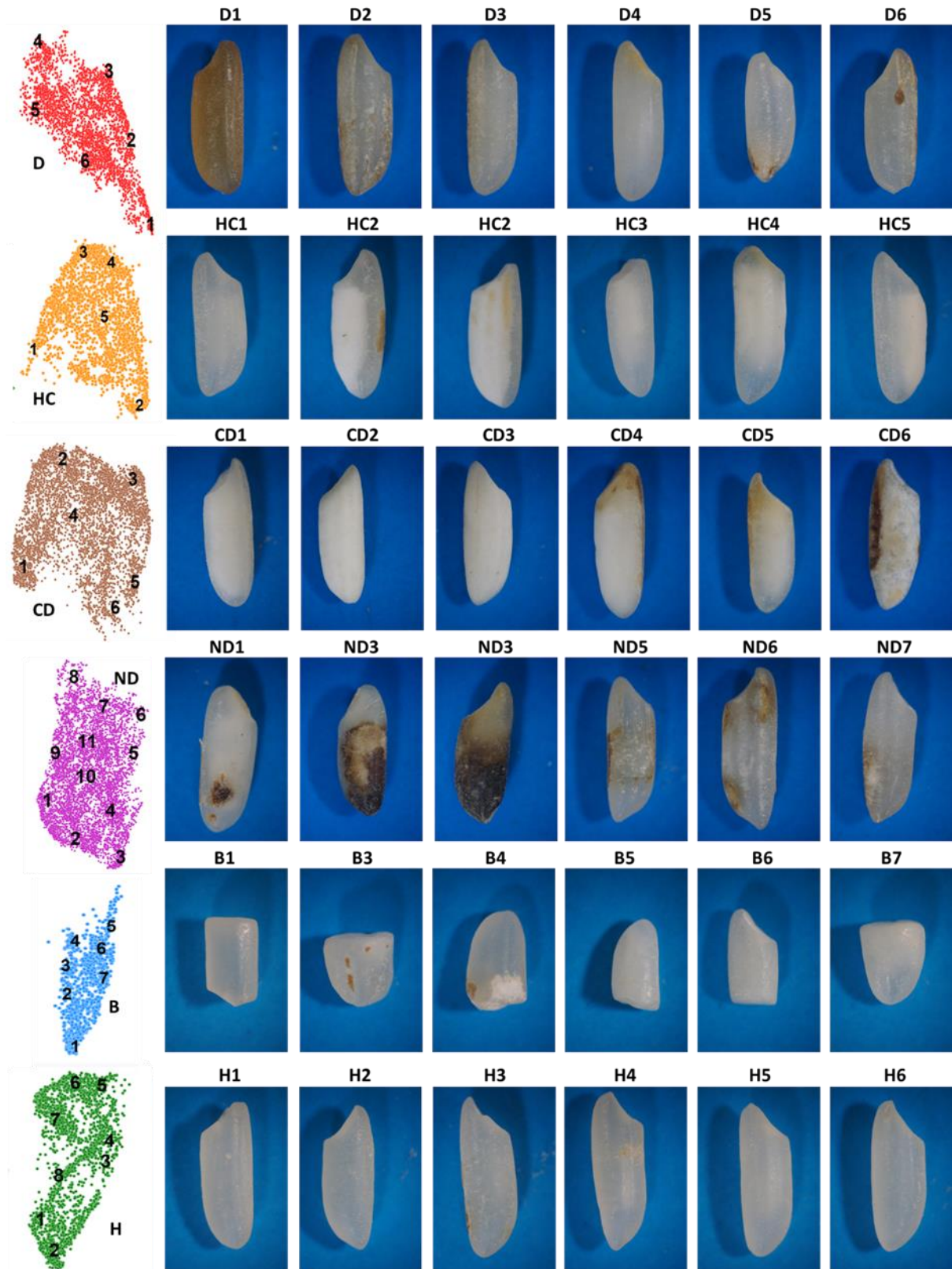
Figure 12. Examples of grains in each cluster spread across different regions. Each row represents a separate class and the corresponding grains. The cluster for each class is separated in different regions marked with numbers. Each image is marked with the class and the region it belongs to. For instance, 'CD4' corresponds to a chalky-discoloured grain belonging to the 4th region as marked in the cluster for chalky-discoloured class.

Starting with the cluster for *discoloured* damage, it is observed that r1 contains highly discoloured grains and as we move from r1 to r4, there is a continuous decrease in the discolouration, with grains in the r4 region being the least discoloured. This was irrespective of the path taken i.e., discolouration decreased along each of the paths 1-2-3-4, 1-7-8-4 and 1-6-5-4. It is also seen that as we move towards r4, chalkiness (very subtle) is being introduced. This is not noticeable near r1 but only becomes noticeable near r8. A considerable number of discoloured grains were also observed to contain a layer of husk. This was evident near r2 and r3, where grains contained a significant layer of husk. Note that these layers of husk were not present on their own but along with discolouration. Grains having very mild layers of husk and no discolouration (or other damage) were a part of the *healthy* class. Further, the grains near r5 and r6 had some characteristics similar to normal-damage; some grains were found with small spots or very mild surface indentations. Notice how r4 is closer to the *chalky-discoloured* cluster (Fig. 8) and is characterised with some chalkiness, and how r5 and r6 are closer to the *normal-damage* cluster (Fig. 8) with similar properties. This is very intuitive and tells how not only the local, but also the global structure is preserved.

In the *half-chalky* class, grains near r1 had a lower degree of chalkiness and shifted more towards the healthy class. Moving towards r3 and r4, the degree of chalkiness increased. Grains near r2 also had a higher degree of chalkiness along with mild layers of husk. In the *chalky-discoloured* cluster, r1, r2 and r3 were predominantly found to contain grains with very high degree of chalkiness and almost no discoloration. Moving lower-right, towards r4, r5 and r6, discolouration started appearing. r5 had a higher degree of discoloration while r6 started having characteristics similar to that of heat and pin damage.

Normal-damage class had the largest number of grains and the most variations. Majority of the grains were found to contain pin damage. Sole watermark and heat damage were comparatively less in number. Pin damage dominated the middle and the lower regions of the cluster. Region r1 contained pin damaged grains with small sized dark indentations. Moving downwards from r1 to r3, grains were found to be more damaged i.e., with darker and larger indentations. Regions r3 contained the grains with the most damage. These were found to be very dark coloured and contained a mix of heat and pin damage. As we moved upwards, the intensity of

the colour started decreasing. Larger sized indentations were still found but with paler colour (r4 and r5). Also, watermark damages started to appear (r5), becoming more prominent around r6. In the upper regions, the right side (r6 and r7) was dominated by watermark damages. The left side (r6) had more pin damages but these were very mild in nature; smaller indentations in very pale in colour. The regions in the middle (r10, r9, r11) were found to contain transitioning properties.

Coming to *broken* damage, the lowermost regions (r1 and r2) were found to contain translucent broken grains without any other damage. On moving upwards, mild discolorations appeared, some in the form of husk (r3). Upper left region r4 had grains with a higher degree of damage with discolouration and indentations. Upper Right regions (r5, r6, r7) had grains with chalkiness, with its degree increasing as we moved upwards. *Healthy* grains had the least number of variations. These grains were found to be very consistent. Similar to broken, the lower regions had the most purity in terms of translucency and absence of any other damage. Leftmost regions (r7, r8) were also found to be very pure. r3 and r4 were found to contain very mild layers of husk or indentations almost not visible to the human eye. The grains in the regions r4 were found to contain more noticeable indentation. The uppermost regions r5 and r6 contained grains with very slight chalkiness.

Thus, the analysis showed how the method contrastive-rc is able to produce semantically meaningful 2D representations with preserving the local as well as the global structure of the image data. The properties of the grains in the various regions of each cluster were explored and it can be seen how intricate the features are separating various regions. These features are very difficult to analyse manually using low dimensional images, thus higher dimensional images were required for manual analysis. However, the method is able to produce such information rich representations with images having a comparatively much lower resolution of 75 x 50 pixels. Further, although the clusters were decently separated, overlaps in damages could still be observed on the borders. For instance, in discoloured damage, the grains in r4 (D4) showed mild chalkiness hence overlapping with the half-chalky and chalky-discoloured class; In chalky-discoloured, grains in r6 (CD6) showed symptoms similar to that of heat damage in the normal-damage class; In half-chalky, grains near r4 (HC4) showed higher

degrees of chalkiness thus overlapping with the higher degree chalky grains in the chalky-discoloured class. Such overlaps are very difficult to identify manually and hence the task of manual labelling becomes very difficult. Further, manual labelling also requires domain knowledge to select the right features and identify the number of classes to separate the grains into. Thus, it highlights how our method is able to tackle all these issues and produce the desirable results.

Although, in each cluster no distinct sub clustering is observed because of the continuous nature of distribution, the different regions can be identified as smaller meaningful groupings showing a well defined set of properties. Many of these regions differ from each other on the basis of degree of damage. Regions in *discoloured* category can be seen to differ on the basis of the amount of discolouration. Regions in *half-chalky* and *chalky-discoloured* differed on the basis of degree of chalkiness. In *normal-damage,* regions were separated on the basis of size and colour. This allows for further subclassification of each damage. Now, since subclassification can be very problem specific and user based, thus there are no hard boundaries proposed for the same. The method contrastive-rc allows for the subclassification to be according to the user's needs and understanding. For any cluster, different regions can be marked according to how the user wants to classify that particular region's grains. For instance, in the *discoloured* cluster, if the user wants grains within a particular amount of discolouration to be counted as "good quality", then those regions can be separately marked as "good quality" under the user's supervision; The other regions will be marked as "poor quality". Further, in *normal-damage,* if a user wants to mark grains with very mild damage (light colour and small size) to be fit for human consumption, then the method allows for the same. The method hence provides a very low-level control for the user allowing a more diverse classification which can be used in multiple scenarios for the classification of rice. This is something which is very difficult to achieve using some classical feature extraction techniques or supervised learning approaches since those algorithms are fundamentally developed to suit a specific task and hence are constrained in that manner.

Chapter 6 - Conclusion and Final Remarks

The following key conclusions can be drawn from this study:

1) A large dataset (20,134 images) is developed of high magnification 24 Megapixel (6000 x 4000 pixels) images of individual rice grains (processed white rice) spread across different damages.

2) A deep unsupervised method Contrastive-RC is developed for fine-grained damage classification of processed white rice, leveraging contrastive self-supervised learning techniques. Extensive experimental analysis was performed to achieve the desired clustering of the 19330 rice grain images, illustrating the importance of the data augmentation strategy and hyperparameter selection.

3) A total of six classes are presented corresponding to each cluster i.e., one Healthy class and five damage classes namely Broken, Half-chalky, Chalky-discoloured, Discoloured, and Normal-damage. Our analysis showed how each of the classes possess distinct well-defined visual properties separating them from each other. Furthermore, it was shown how each type differs in quality and hence in market value, thus showcasing the need for their separation.

4) It is shown how the method Contrastive-RC is able to produce meaningful and information rich 2D embeddings from comparatively lower resolution (75 x 50) images. Our analysis showed that the method is robust towards variations in brightness and grain geometry (size, orientation, aspect ratio, rotation), thus making its extension to other varieties very feasible. The method is also fast with the ability to process 370 grain images per second making it feasible for real-world use cases where lower latency is required.

5) An extensive analysis of the different regions in each cluster showed that the method is able to preserve the local as well as the global structure of the image data. It was shown how different regions inside each cluster can be identified as smaller meaningful groupings, often separated on the basis of the degree of damage. And

how this allows for a user-supervised subclassification, which can be used in multiple use cases, providing very low level control for various tasks.

The method Contrastive-RC not only facilitates the broader classification of rice grain images into six damage-based classes but also provides a means for further subclassification providing low level control allowing for it to be used in multiple use cases.

The future scope of this project would be to extend the algorithm to be able to work on images taken by mobile cameras as a proof of concept. Further, to design an end-to-end machine that can leverage the computer vision algorithm developed to provide real time quality evaluation and sorting of food grains based on various types.

References

[1] Grunert, K. G. (2005). Food quality and safety: consumer perception and demand. *European review of agricultural economics*, *32*(3), 369-391.

[2] Mascarello, G., Pinto, A., Parise, N., Crovato, S., & Ravarotto, L. (2015). The perception of food quality. Profiling Italian consumers. *Appetite*, *89*, 175-182.

[3] "Top ten India crops and agricultural products," can be found under https://www.mapsofindia.com/top-ten/india-crops/, 2021.

[4] *Statista - The Statistics Portal*. (n.d.). Statista; www.statista.com. Retrieved May 24, 2022, from https://www.statista.com

[5] Vithu, P., & Moses, J. A. (2016). Machine vision system for food grain quality evaluation: A review. *Trends in Food Science & Technology*, *56*, 13-20. [6] Custodio, M. C., Cuevas, R. P., Ynion, J., Laborte, A. G., Velasco, M. L., & Demont, M. (2019). Rice quality: How is it defined by consumers, industry, food scientists, and geneticists?. *Trends in food science & technology*, *92*, 122-137.

[7] Tashiro, T. and Wardlaw, I., 2022. *The effect of high temperature on kernel dimensions and the type and occurrence of kernel damage in rice*.

[8] Morita, S., Wada, H., & Matsue, Y. (2016). Countermeasures for heat damage in rice grain quality under climate change. *Plant Production Science*, *19*(1), 1-11.

[9] Chen, Z., Wassgren, C., & Ambrose, K. (2020). A review of grain kernel damage: mechanisms, modeling, and testing procedures. *Transactions of the ASABE*, *63*(2), 455-475.

[10] Sugimoto, A., & Nugaliyadde, L. (1995). Damage of rice grains caused by the rice bug, Leptocorisa oratorius (Fabricius)(Heteroptera: Alydidae). *Jpn. Int. Res. Center Agric. Sci. J*, *2*, 13-17.

[11] Nakata, M., Fukamatsu, Y., Miyashita, T., Hakata, M., Kimura, R., Nakata, Y., ... & Yamakawa, H. (2017). High temperature-induced expression of rice α-amylases in developing endosperm produces chalky grains. *Frontiers in plant science*, *8*, 2089.

[12] Y.-N. Wan, et al. 'RICE QUALITY CLASSIFICATION USING AN AUTOMATIC GRAIN QUALITY INSPECTION SYSTEM'. Transactions of the ASAE, vol. 45, no. 2, 2002. DOI.org (Crossref), https://doi.org/10.13031/2013.8509.

[13] Fitzgerald, M. A., McCouch, S. R., & Hall, R. D. (2009). Not just a grain of rice: the quest for quality. *Trends in plant science*, *14*(3), 133-139.

[14] https://fci.gov.in/

[15] Graham, R. (2002). *A proposal for IRRI to establish a grain quality and nutrition research center* (No. 2169-2019-1615).

[16] Mihajlovski, Katarina, et al. 'Valorization of Damaged Rice Grains: Optimization of Bioethanol Production by Waste Brewer's Yeast Using an Amylolytic Potential from the Paenibacillus Chitinolyticus CKS1'. Fuel, vol. 224, July 2018, pp. 591–99. DOI.org (Crossref), https://doi.org/10.1016/j.fuel.2018.03.135.

[17] Paliwal, J., Borhan, M. S., & Jayas, D. S. (2003). Classification of cereal grains using a flatbed scanner. In *2003 ASAE Annual Meeting* (p. 1). American Society of Agricultural and Biological Engineers.

[18] Pearson, T. (2010). High-speed sorting of grains by color and surface texture. *Applied engineering in agriculture*, *26*(3), 499-505.

[19] Patel, K. K., Kar, A., Jha, S. N., & Khan, M. A. (2012). Machine vision system: a tool for quality inspection of food and agricultural products. *Journal of food science and technology*, *49*(2), 123-141.

[20] El-Mesery, H. S., Mao, H., & Abomohra, A. E. F. (2019). Applications of non-destructive technologies for agricultural and food products quality inspection. *Sensors*, *19*(4), 846.

[21] Rehman, T. U., Mahmud, M. S., Chang, Y. K., Jin, J., & Shin, J. (2019). Current and future applications of statistical machine learning algorithms for agricultural machine vision systems. *Computers and electronics in agriculture*, *156*, 585-605.

[22] Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and electronics in agriculture*, *147*, 70-90.

[23] Magomadov, V. S. (2019, December). Deep learning and its role in smart agriculture. In *Journal of Physics: Conference Series* (Vol. 1399, No. 4, p. 044109). IOP Publishing.

[24] Mery, D., Pedreschi, F., & Soto, A. (2013). Automated design of a computer vision system for visual food quality evaluation. *Food and Bioprocess Technology*, *6*(8), 2093-2108.

[25] Kaur, H., & Singh, B. (2013). Classification and grading rice using multi-class SVM. *International Journal of Scientific and Research Publications*, *3*(4), 1-5.

[26] Aggarwal, A. K., & Mohan, R. (2010). Aspect ratio analysis using image processing for rice grain quality. *International Journal of Food Engineering*, *6*(5).

[27] Van Dalen, G. (2004). Determination of the size distribution and percentage of broken kernels of rice using flatbed scanning and image analysis. *Food research international*, *37*(1), 51-58.

[28] Guzman, J. D., & Peralta, E. K. (2008, August). Classification of Philippine rice grains using machine vision and artificial neural networks. In *World conference on Agricultural information and IT* (Vol. 6, pp. 41-48).

[29] Lee, C. Y., Yan, L., Wang, T., Lee, S. R., & Park, C. W. (2011). Intelligent classification methods of grain kernels using computer vision analysis. *Measurement science and Technology*, *22*(6), 064006.

[30] Ozan, A. K. I., Güllü, A., & Uçar, E. (2015, November). Classification of rice grains using image processing and machine learning techniques. In *International scientific conference* (pp. 20-21).

[31] Chen, S., Xiong, J., Guo, W., Bu, R., Zheng, Z., Chen, Y., ... & Lin, R. (2019). Colored rice quality inspection system using machine vision. *Journal of cereal science*, *88*, 87-95.

[32] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, *521*(7553), 436-444.

[33] Li, Z., Liu, F., Yang, W., Peng, S., & Zhou, J. (2021). A survey of convolutional neural networks: analysis, applications, and prospects. *IEEE Transactions on Neural Networks and Learning Systems*.

[34] Wu, Q., Liu, Y., Li, Q., Jin, S., & Li, F. (2017, October). The application of deep learning in computer vision. In *2017 Chinese Automation Congress (CAC)* (pp. 6522-6527). IEEE.

[35] Voulodimos, A., Doulamis, N., Doulamis, A., & Protopapadakis, E. (2018). Deep learning for computer vision: A brief review. *Computational intelligence and neuroscience*, *2018*.

[36] Ibrahim, S., Kamaruddin, S. B. A., Zabidi, A., & Ghani, N. A. M. (2020). Contrastive analysis of rice grain classification techniques: multi-class support vector machine vs artificial neural network. *IAES International Journal of Artificial Intelligence*, *9*(4), 616.

[37] Lin, P., Chen, Y., He, J., & Fu, X. (2017, December). Determination of the varieties of rice kernels based on machine vision and deep learning technology. In *2017 10th International Symposium on Computational Intelligence and Design (ISCID)* (Vol. 1, pp. 169-172). IEEE.

[38] Marini, F., Zupan, J., & Magrì, A. L. (2004). On the use of counterpropagation artificial neural networks to characterize Italian rice varieties. *Analytica chimica acta*, *510*(2), 231-240.

[39] Aukkapinyo, K., Sawangwong, S., Pooyoi, P., & Kusakunniran, W. (2020). Localization and classification of rice-grain images using region proposals-based convolutional neural network. *International Journal of Automation and Computing*, *17*(2), 233-246.

[40] Chatnuntawech, I., Tantisantisom, K., Khanchaitit, P., Boonkoom, T., Bilgic, B., & Chuangsuwanich, E. (2018). Rice classification using spatio-spectral deep convolutional neural network. *arXiv preprint arXiv:1805.11491*.

[41] Payman, S. H., Bakhshipour, A., & Zareiforoush, H. (2018). Development of an expert vision-based system for inspecting rice quality indices. *Quality Assurance and Safety of Crops & Foods*, *10*(1), 103-114.

[42] Wan, Y. N., Lin, C. M., & Chiou, J. F. (2002). Rice quality classification using an automatic grain quality inspection system. *Transactions of the ASAE*, *45*(2), 379.

[43] Ruan, R., Ning, S., Song, A., Ning, A., Jones, R., & Chen, P. (1998). Estimation of Fusarium scab in wheat using machine vision and a neural network. *Cereal chemistry*, *75*(4), 455-459.

[44] Jirsa, O., & Polišenská, I. (2011). Identification of Fusarium damaged wheat kernels using image analysis. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, *59*(5), 125-130.

[45] Wiwart, M., Koczowska, I., & Borusiewicz, A. (2001, September). Estimation of Fusarium head blight of triticale using digital image analysis of grain. In *International Conference on Computer Analysis of Images and Patterns* (pp. 563-569). Springer, Berlin, Heidelberg.

[46] Steenhoek, L. W., Misra, M. K., Hurburgh Jr, C. R., & Bern, C. J. (2001). Implementing a computer vision system for corn kernel damage evaluation. *Applied Engineering in Agriculture*, *17*(2), 235.

[47] Pazoki, A. R., Farokhi, F., & Pazoki, Z. (2014). Classification of rice grain varieties using two Artificial Neural Networks (MLP and Neuro-Fuzzy). *The Journal of Animal & Plant Sciences*, *24*(1), 336-343.

[48] Lin, P., Li, X. L., Chen, Y. M., & He, Y. (2018). A deep convolutional neural network architecture for boosting image discrimination accuracy of rice species. *Food and Bioprocess Technology*, *11*(4), 765-773.

[49] Lizhang, X., & Yaoming, L. (2008, September). Multi-scale edge detection of rice internal damage based on computer vision. In *2008 IEEE International Conference on Automation and Logistics* (pp. 1222-1225). IEEE.

[50] Wang, D., Dowell, F. E., Lan, Y., Pasikatan, M., & Maghirang, E. (2002). Determining pecky rice kernels using visible and near-infrared spectroscopy. *International Journal of Food Properties*, *5*(3), 629-639.

[51] Aggarwal, C. C., & Reddy, C. K. (2014). Data clustering. *Algorithms and applications. Chapman&Hall/CRC Data mining and Knowledge Discovery series, Londra*.

[52] K. Gowda and G. Krishna. Agglomerative clustering using the concept of mutual nearest neighbourhood. Pattern Recognition, 10(2):105–112, 1978.

[53] R. Vidal. Subspace clustering. IEEE Signal Processing Magazine, 28(2):52–68, 2011.

[54] J. Wang, J. Wang, J. Song, X. Xu, H. Shen, and S. Li. Optimized cartesian k-means. IEEE Trans. Knowl. Data Eng., 27(1):180–192, 2015.

[55] Y. Yang, D. Xu, F. Nie, S. Yan, and Y. Zhuang. Image clustering using local discriminant models and global integration. T-IP, 19(10):2761–2773, 2010.

[56] J. Yu. General c-means clustering model and its application. In CVPR, pages 122–127, 2003.

[57] Steinbach, M., Ertöz, L., & Kumar, V. (2004). The challenges of clustering high dimensional data. In *New directions in statistical physics* (pp. 273-309). Springer, Berlin, Heidelberg.

[58] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In CVPR, pages 886–893, 2005.

[59] D. Lowe. Distinctive image features from scale-invariant keypoints. IJCV, 60(2):91–110, 2004

[60] Chen, T., Kornblith, S., Norouzi, M., & Hinton, G. (2020, November). A simple framework for contrastive learning of visual representations. In *International conference on machine learning* (pp. 1597-1607). PMLR.

[61] Chen, X., & He, K. (2021). Exploring simple siamese representation learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 15750-15758).

[62] Hu, H., Bindu, J. P., & Laskin, J. (2022). Self-supervised clustering of mass spectrometry imaging data using contrastive learning. *Chemical science*, *13*(1), 90-98.

[63] Ciortan, M., & Defrance, M. (2021). Contrastive self-supervised clustering of scRNA-seq data. *BMC bioinformatics*, *22*(1), 1-27.

[64] Ayush, K., Uzkent, B., Meng, C., Tanmay, K., Burke, M., Lobell, D., & Ermon, S. (2021). Geography-aware self-supervised learning. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 10181-10190).

[65] McInnes, L., Healy, J., & Astels, S. (2017). hdbscan: Hierarchical density based clustering. *J. Open Source Softw.*, *2*(11), 205.

[66] Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern recognition letters*, *31*(8), 651-666.

[67] Steinbach, M., Ertöz, L., & Kumar, V. (2004). The challenges of clustering high dimensional data. In *New directions in statistical physics* (pp. 273-309). Springer, Berlin, Heidelberg.

[68] Coleman, G. B., & Andrews, H. C. (1979). Image segmentation by clustering. *Proceedings of the IEEE*, *67*(5), 773-785.

[69] Wazarkar, S., & Keshavamurthy, B. N. (2018). A survey on image data analysis through clustering techniques for real world applications. *Journal of Visual Communication and Image Representation*, *55*, 596-626.

[70] Jaiswal, A., Babu, A. R., Zadeh, M. Z., Banerjee, D., & Makedon, F. (2020). A survey on contrastive self-supervised learning. *Technologies*, *9*(1), 2.

[71] Van Gansbeke, W., Vandenhende, S., Georgoulis, S., Proesmans, M., & Van Gool, L. (2020, August). Scan: Learning to classify images without labels. In *European Conference on Computer Vision* (pp. 268-285). Springer, Cham.

[72] Tian, Y., Krishnan, D., & Isola, P. (2020, August). Contrastive multiview coding. In *European conference on computer vision* (pp. 776-794). Springer, Cham.

[73] Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of big data*, *6*(1), 1-48.

[74] Howard, A., Sandler, M., Chu, G., Chen, L. C., Chen, B., Tan, M., ... & Adam, H. (2019). Searching for mobilenetv3. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 1314-1324).

[75] Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

[76] 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.

[77] McInnes, L., Healy, J., & Melville, J. (2018). Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*.

[78] Van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of machine learning research*, *9*(11).

[79] https://www.ams.usda.gov/sites/default/files/media/VRI2017.pdf

[80] Keras: The Python Deep Learning API. https://keras.io/.

[81] TensorFlow, https://www.tensorflow.org/.

[82] UMAP, https://umap-learn.readthedocs.io/en/latest/

[83] HDBSCAN, https://hdbscan.readthedocs.io/en/latest/index.html

[84] sklearn, https://hdbscan.readthedocs.io/en/latest/index.html

[85] Siebenmorgen, T. J., Counce, P. A., & Wilson, C. E. (2011). Factors affecting rice milling quality.

[86] Demont, M., Fiamohe, R., & Kinkpe, A. T. (2017). Comparative advantage in demand and the development of rice value chains in West Africa. *World Development*, *96*, 578-590.

[87] Adair, C. R. (1966). Rice in the United States: varieties and production. *Rice in the United States: varieties and production.*

[88] Graham, R. (2002). *A proposal for IRRI to establish a grain quality and nutrition research center* (No. 2169-2019-1615).

[89] Unnevehr, L., Duff, B., & Juliano, B. O. (Eds.). (1992). *Consumer demand for rice grain quality: terminal report of IDRC projects, National grain quality (Asia), and International grain quality economics (Asia)*. Int. Rice Res. Inst..

[90] Zhao X, Fitzgerald MA (2013) Climate change: implications for the yield of edible rice. PLoS One 8(6):e66218

[91] https://www.ams.usda.gov/sites/default/files/media/VRI2017.pdf

[92] https://fci.gov.in/

[93] Ndindeng, S. A., Candia, A., Mapiemfu, D. L., Rakotomalala, V., Danbaba, N., Kulwa, K., ... & Futakuchi, K. (2021). Valuation of rice postharvest losses in sub-Saharan Africa and its mitigation strategies. *Rice Science*, *28*(3), 212-216.

[94] Brorsen, B. W., Grant, W. R., & Rister, M. E. (1984). A hedonic price model for rough rice bid/acceptance markets. *American Journal of Agricultural Economics*, *66*(2), 156-163.

[95] United States Department of Agriculture. 2009. United States standards for rice. Accessed July 10, 2020. https://www. gipsa.usda.gov/fgis/standards/ricestandards.pdf

[96] Baiardi D., Puglisi R., Scabrosetti S. Individual attitudes on food quality and safety: Empirical evidence on EU countries. Food Qual. Prefer. 2016;49:70–74. doi: 10.1016/j.foodqual.2015.11.011

[97] https://scikit-learn.org/stable/modules/clustering.html

[98] Chun, A., Song, J., Kim, K. J., & Lee, H. J. (2009). Quality of head and chalky rice and deterioration of eating quality by chalky rice. *Journal of Crop Science and Biotechnology*, *12*(4), 239-244.

[99] Brorsen, B. W., Grant, W. R., & Rister, M. E. (1984). A hedonic price model for rough rice bid/acceptance markets. *American Journal of Agricultural Economics*, *66*(2), 156-163.

[100] UMAP, https://github.com/lmcinnes/umap

[101] Bhupendra, Moses K., Miglani, A., & Kankar, P. K. (2022). Deep CNN-based damage classification of milled rice grains using a high-magnification image dataset. *Computers and Electronics in Agriculture*, *195*, 106811.

[102] Unnevehr, Laurian, et al., editors. Consumer Demand for Rice Grain Quality: Terminal Report of IDRC Projects, National Grain Quality (Asia), and International Grain Quality Economics (Asia). International Rice Research Institute, 1992.

[103] Brorsen, B. Wade, et al. 'A Hedonic Price Model for Rough Rice Bid/Acceptance Markets'. American Journal of Agricultural Economics, vol. 66, no. 2, May 1984, pp. 156–63. DOI.org (Crossref), https://doi.org/10.2307/1241032