



DEPARTMENT OF CHEMISTRY AND MOLECULAR
BIOLOGY

Correlation of Visual Cues with In Vitro Protein Digestion using Computer Vision

Julia Mårtensson

Essay/Thesis:	30 hp
Program and/or course:	Bachelor's Thesis in Chemistry, KEM929
Level:	First Cycle
Term/year:	Summer-Autumn 2024
Supervisors:	Giovanni Tizzanini, Jakob Ytterberg and Anna Ström
Examiner:	Leif Eriksson

Abstract

Digestion studies are crucial in medical research and diagnostics, yet existing *in vitro* models lack realism, while *in vivo* methods are often invasive, risky, and costly. Although endoscopy has become widely used in gastrointestinal studies, it frequently requires complementary biopsies and microbiota sampling to provide a complete picture. This thesis explores the feasibility of using capsule endoscopy to observe and quantitatively measure digestion directly from the camera data, through the development and evaluation of a prototype *in vitro* protocol and computer vision. In this setup, a similar small camera, coated in gelatin hydrogel is immersed in media simulating gastrointestinal conditions, capturing video as the colored gel dissolves. Techniques like swelling tests and digested media analysis can help pinpoint important highlights of protein digestion. Computer vision analysis can successfully identify the point of full gel dissolution, but quantifying the digestion rate before this point proved challenging due to color inaccuracies in the USB camera microscope, which post-processing could not correct. Future research should consider color-accurate cameras or protocols less dependent on color, such as tracking the smearing of a patterned gel. Finally, generalizing from an *in vitro*, ambient lit setup to the internally lit environment of capsule endoscopy may unlock new techniques for measuring the digestion, requiring further study.

story dominated by the realization that color-accurate work require color-accurate cameras but also raising questions about color-accuracy-resistant alternatives.

Essay/Thesis:	30 hp
Program and/or course:	Bachelor's Thesis in Chemistry, KEM929
Level:	First Cycle
Term/year:	Summer-Autumn 2024
Supervisors:	Giovanni Tizzanini, Jakob Ytterberg and Anna Ström
Examiner:	Leif Eriksson
Keywords:	Gelatin Hydrogels, Gel Digestibility and Dissolution, Digestion Monitoring, Nutrition, Computer Vision, Image Analysis

Preface

This thesis is part of the PAN Sweden project, specifically the Precision Nutrition Innovation arena (PNI). PAN Sweden is a research collaboration center between Örebro University, Chalmers University of Technology, Uppsala University, the Swedish University of Agricultural Sciences (SLU), and the Research Institutes of Sweden (RISE). This Bachelor's thesis is part of a project within the PAN Sweden project where the overarching goal is to develop a proof of concept demonstrating the feasibility of non-invasive *in vivo* methods, such as capsule endoscopy, for observing and measuring the digestion of food matter directly.

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Giovanni Tizzanini, for his invaluable support and continuous encouragement throughout this project. Our meetings consistently left me with renewed confidence and discussing ideas about the project with him has both been very interesting and insightful. His understanding and guidance have been instrumental to my progress and I am truly grateful to have had him as my supervisor.

I also wish to thank my co-supervisors, Jakob Ytterberg and Anna Ström, for their valuable feedback, help, and support. I've much appreciated Jakob's help with 3D printing and both Anna and Jakob's insightful input and comments throughout this project. I would also like to thank Martin Längkvist from Örebro University for the great discussions regarding the computational parts of this project. Additionally, I would like to thank Leif Eriksson, my examiner in this project.

Finally, I am thankful to have had the opportunity to work with food chemistry and computer vision in the Applied Chemistry division at Chalmers University of Technology, and for having been a part of the PAN Sweden and PNI project. I'm very thankful to everyone for making me feel warmly welcomed and included at the division. I've truly had a great time!

Contents

1	Introduction	4
1.1	Digestion	4
1.2	Goal of this Thesis	4
1.3	Computer Vision	5
1.4	Linear Regression	5
1.5	Canny Edge Detection	5
1.6	Hough Circle Detection	5
1.7	Color Spaces	6
1.8	Gels and Hydrogels	6
2	Materials and Method	7
2.1	Chemicals	7
2.2	Gelatin Gel Casting	7
2.3	Compression Tests	7
2.4	Media Composition	8
2.5	Swelling Test	9
2.6	UV/Vis and Protein Loss Determination	10
2.7	Video Recordings	10
2.8	Image Analysis	10
3	Results and Discussion	11
3.1	Gel Characterization	11
3.1.1	Compression Test	12
3.1.2	Swelling Tests	12
3.1.3	UV/Vis Spectroscopy	13
3.2	Steps of Image Analysis	15
3.2.1	Clip to Middle Square	16
3.2.2	Brightness vs. Saturation	16
3.2.3	Bubble Removal Using Hough Circle Detection	17
3.2.4	Discontinuity Removal	17
3.2.5	Linear Regression	17
3.3	Gel Digestion and Protocol Evaluation	18
4	Conclusion	23
References		24

1 Introduction

In the following sections, a theoretical overview is given of the digestion process, limitations of existing *in vitro* and *in vivo* models, and the potential of computer vision in advancing these studies. Additionally, the theory introduces a brief overview of gels, hydrogels, gelation, and crosslinking, since crosslinked gelatin hydrogels are used to observe and measure how digestion and dissolution processes occur.

1.1 Digestion

Digestion, fermentation, and adsorption of food, drugs, and other edible materials is a relevant field. From nutrition to drug delivery, from medical conditions that affect the gastrointestinal tract to specific ingredients digestibility assessments. For instance, globally there are billions of cases and millions of deaths due to digestive diseases every year [1], drawing billions in research funds every year [2]. Different models are employed in industry and academia to study the digestibility of edible matrices. These models can be *in vivo*, where the study is conducted on a living organism, either human or animal, or *in vitro*, where lab-based models are used to simulate parts of the gastrointestinal (GI) tract.[3]

Examples of *in vivo* techniques might be carried out on animal models, such as rodents, piglets, or guinea pigs, or humans; they consist of biopsy sampling or intubation techniques, which, for humans, are often carried out on patients with pre-existing pathologies that require intubation, like enterostomy or fistulas. Breath and feces analysis are other common analyses to carry out studies on digestion and metabolism. *In vitro* techniques might be static, as the INFOGEST protocol and the human colonic fermentation, or dynamic, as the TNO Gastrointestinal Model (TIM-1, TIM-2), where a dynamic, computer-controlled system mimics the GI tract.[4]

Being the only available and widely recognized models, they are extensively employed, but they all show various limitations on their use and applicability. *In vivo* models have clear ethical limits, for invasive, uncomfortable, and risky procedures on the patients, and, besides, deadly for animals. The patients' number is limited, the allowed age is restricted, and the tests are costly and time-consuming, requiring extensive preparation, ethical approval, and specialized equipment. Furthermore, animal models can provide insights into the human GI, but the species differences limit the direct applicability to humans, and patients suffering enterostomy or fistulas have an altered and non-representative digestive process. Static *in vitro* systems cannot fully replicate how dynamic the GI tract is (e.g.: gut motility and pH, enzymatic and salt gradients), and there is no interconnection between the different modeled GI tracts. For microbiota representation, the techniques average everything and fail to accurately replicate the *in vivo* microbial succession and spatial distribution along the colon. On the other hand, dynamic techniques are high-end, expensive, and with limited access, therefore hard to employ widely in the field. For both, the conditions are oversimplified, and the complex host biology is absent here. For this reason, nutritional and pharmaceutical interventions are often tested in *in vitro* models first to narrow down candidates, before moving to more costly and complex *in vivo* studies.[5] [6] [7]

1.2 Goal of this Thesis

The goal of this thesis is to produce an *in vitro* proof of concept for using computer vision to measure the digestion of a TGase-crosslinked gelatin hydrogel immersed into gastrointestinal-simulating media and to predict the properties of the media based on the rate of digestion. Properties varying in the media are pH, salinity, and the presence of pepsin, a protein-digesting enzyme. The aim is to investigate whether computer vision techniques can accurately and reliably determine the chemical configuration of the media based on visual cues of recordings taken at the digesting media. Therefore, a bio-compatible and edible gelatin gel

will be cast to cover a camera submerged in different *in vitro* media simulating stomach juice. To enhance the technical properties and stability of the gel, transglutaminase (TGase) is used to create enzyme-induced crosslinks, raising the melting point of the gel beyond the physiological one of 37°C. This bachelor's thesis is encased in the PNI project and will build the necessary dataset and computer vision using a similar microscope probe camera and an *in vitro* system.

1.3 Computer Vision

Computer vision is a field focused on enabling computers to interpret and analyze visual data from images and videos. It involves techniques for processing, identifying, and extracting information from visual data. This enables computers to recognize objects and make quantitative assessments of visual data [8]. Applications of computer vision cover a vast range of fields. For example, computer vision can be used in medical imaging to analyze X-rays or MRI scans, helping detect abnormalities like tumors, as well as being used for detecting anomalies in fruits, food, and food sorting [9]. Computer vision has three main advantages over human judgment. It is automated, rather than manual labor, it is quantitative, and it can pick up on more subtle cues by leveraging large amounts of training data.

1.4 Linear Regression

Linear regression aims to find the line that fits a dataset the best by minimizing variance, specifically by minimizing the sum of squared errors (SSE). Given a line $y = \beta_0 + \beta_1 x$, where β_0 and β_1 are parameters, the SSE is defined as,

$$S(\beta_0, \beta_1) = \sum (y_i - (\beta_0 + \beta_1 x_i))^2,$$

which is proportional to the variance of the dataset after removing the trend line. Since S is quadratic in β_0 and β_1 , partial derivatives are taken with respect to each parameter, set to zero, and the resulting linear system is solved. This process yields values for β_0 and β_1 that minimize the SSE, thereby determining the best-fit line. [10]

1.5 Canny Edge Detection

Canny edge detection [11] is an algorithm that identifies edges by first reducing noise with a Gaussian blur, and then calculating intensity gradients to detect edge areas. It refines edges by keeping only peak gradient values, applies thresholds to categorize edges as strong or weak, and finally connects weak edges to strong ones while discarding isolated points. Canny edge detection effectively isolates true edges, minimizes false detections, and is not very sensitive to noise. It is implemented in OpenCV [12].

1.6 Hough Circle Detection

Hough circle detection [13] is used to identify bubbles in the recordings. First, the contrast of each frame is increased to make features more visible. Edges in the high-contrast image are then detected using Canny edge detection, producing a set of edge points. Based on a predetermined range of possible radii, different circle sizes are tested by replacing each detected edge point with a circle of radius r , centered on the point itself. Hough circle detection then identifies locations in the frame with the highest density of circle intersections, which correspond to the centers of circles in the original frame. Hough circle detection is implemented in OpenCV [12].

Figure 1 depicts two different instances of a frame with increased contrast, edge detection, and a mask with Hough circles. The high contrast frame is normalized so that its darkest pixels are fully black and its whitest pixels are fully white.

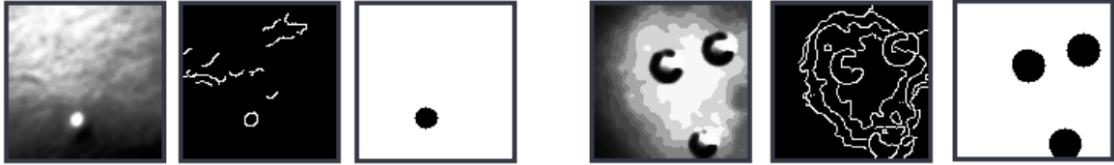


Figure 1: Bubble detection of two different recordings. In both of the recordings bubbles appear and are detected through the Hough circle detection algorithm, the first step in which is edge detection.

1.7 Color Spaces

Color spaces serve as coordinate systems for representing real-world colors in a structured, digital format, even though color spaces themselves do not exist naturally. They map the infinite variety of real-world colors into defined parameters, allowing consistent representation, manipulation, and storage in digital systems.

Some color spaces use linear mapping to represent color intensity directly proportional to light intensity, such as raw RGB data. However, others, like sRGB, apply a power scale called gamma correction to better reflect human perception [14]. This adjustment accounts for the fact that human vision is roughly logarithmic, as humans are more sensitive to differences in darker tones than in lighter ones [15]. By encoding colors in a way that aligns with this sensitivity, color spaces like sRGB achieves perceptual uniformity in luminosity and dedicates higher color resolution to bands where it is relevant.

The HSV color space organizes colors in a way that aligns with human perception by using three components: Hue, representing the color type on a circular scale; Saturation, indicating color intensity; and Value, for brightness [16]. Unlike RGB’s direct mapping of light intensities, HSV allows for intuitive color manipulation by separating tone, intensity, and brightness, making it especially useful in image editing and computer vision applications to measure color along these axes.

1.8 Gels and Hydrogels

Gels are highly solvated three-dimensional networks of cross-linked polymers. Gels are mostly liquid by mass, yet they behave like solids because of a three-dimensional cross-linked network within the liquid. They exhibit no solvent flow in the steady state, although the liquid phase may still diffuse through it. For these characteristics, they find a multitude of applications ranging from biomedical to environmental and industrial uses [17]. Hydrogels are a type of gel where the solvent is water and the polymers are hydrophilic. The hydrophilic nature of the polymers enables the hydrogels to absorb and retain large quantities of water [18].

The name “gel” comes from “gelatin”, which, in turn, comes from the Latin verb *gelare* which translates to “to freeze” [19]. Gelatin is a protein derived from collagen, the main structural protein in the extracellular matrix of an animal’s various connective tissues, and has long been known to humans. Gelatin forms a heat-induced hydrogel when dissolved in warm water and cooled, and it is widely used in various biomedical and industrial applications [20].

The gelation mechanisms stem from the crosslinking of the polymers that constitute the solid part. Crosslinking in gels can occur through chemical reactions, forming chemical gels with covalent bonds, or through physical interactions, resulting in physical gels with non-covalent bonds [21]. The gelation

mechanisms can be heat-induced, pressure-induced, ion-induced, acid-induced, or enzyme-induced [22]. Crosslinking creates a three-dimensional structure that traps and retains the solvent within the network [21] [23] [24]. Transglutaminase (TGase) can be used to create enzyme-induced crosslinks, and in this project, it will be employed to prevent the dissolution of the gelatin gel at physiological temperature.

2 Materials and Method

The method can be divided into three main parts: gel characterization, video recording, and image analysis. Gel characterization was conducted using compression and swelling tests, along with UV/Vis spectroscopy on media from swelling tests with pepsin. Video recordings were performed primarily using a USB camera microscope coated with TGase-crosslinked gelatin hydrogel, placed in media of varying pH, salinity, and pepsin content. In image analysis, quantitative data were extracted from the recordings through steps including clipping to the middle square, masking air bubbles, calculating average HSV saturation, and filtering out discontinuities, culminating in time-series plots with overlaid linear regression.

2.1 Chemicals

The chemicals employed in this project are gelatin type A (Bloom 300 g, porcine skin, Sigma-Aldrich); ultrapure ion free Milli-Q (Millipore, 18.2 MΩ cm resistivity) water; TGase (Galaya® Prime, Novozyme); yellow food coloring dye (Dr. Oetker™); sodium chloride (99% purity, Acros-Organics); hydrochloric acid (2M, Sigma-Aldrich); calcium chloride dihydrate (Sigma-Aldrich).

2.2 Gelatin Gel Casting

The gel composition and casting technique were held constant throughout the project. Gelatin powder (3% m/V) was left swelling at room temperature in Milli-Q water for 30 minutes under continuous stirring. After raising the temperature to 45°C, TGase was added (433 U/g of gelatin in a solution, 13 U/ml Milli-Q water) and let act for 40 min under continuous stirring. Yellow food dye was added to the solution under continuous stirring as well. TGase was deactivated for 5 min in an 80°C water bath under continuous stirring. When cooling down, the gel was cast in the specific molds, see Figure 2, and let set overnight. The food dye was employed to make the gel not completely transparent. The dye had to be edible and the yellow was selected, specifically, because it bonded very well to gelatin and did not leech into solution when immersed in water.

The casting of the USB camera microscope (IP67, Staright) was carried out by pipetting gel into the tip (see Figure 3) of the USB camera microscope. The tip before the video sensor has a cylindrical form and is 2 mm deep and 8 mm wide. To minimize air bubbles, a 3D-printed cylindrical mold was lined with gel so that no air bubbles were formed when the tip of the USB camera microscope was pressed to the plastic base plate. The gel was cast as in Figure 2a.

The camera pill (CapsoCam ®, Gothia Medical) was cast by TGase crosslinked gelatin hydrogel being pipetted into three of the four different sections of the gel mold, see Figure 4. The fourth section was not filled with anything at all in order to get a reference while recording.

2.3 Compression Tests

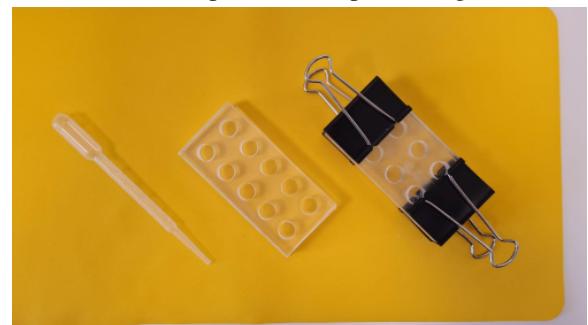
Compression tests of gelatin gels crosslinked with or without the help of TGase were performed on an Instron 68TM-5 testing machine with a 5 kN load cell interfaced with Bluehill 2 software. The gel composition was fixed at 3% m/V, and a deformation rate of 0.2 mm/s was employed. Seven repetitions with cylindrical gelatin gels were made for each sample, all at room temperature.



(a) Setup for USB camera microscope gel casting.



(b) Setup for camera pill casting.



(c) Test molds for swelling test and compression test.

Figure 2: Gel casting for different use cases.

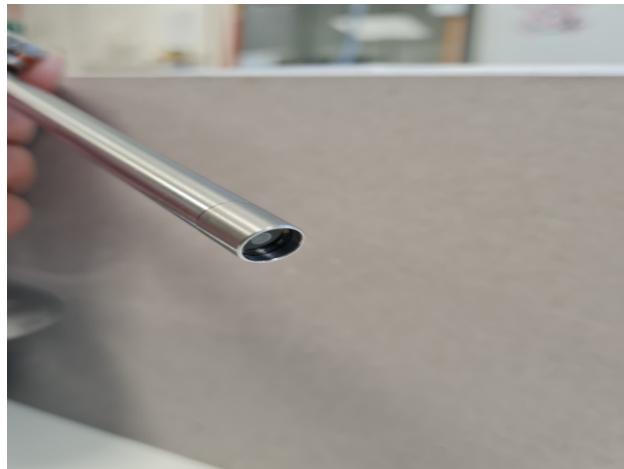


Figure 3: The USB camera microscope tip that the gel is pipetted into. The microscope with the gel cast is set overnight and used to record videos in every media.

2.4 Media Composition

As the gelatin composition was kept constant, the media in which the gels were immersed, changed. The objective was to apply CV in physiological range values that are present in our GI tract and train it to discern



Figure 4: Camera pill where 3 out of 4 sections are filled with gel. The fourth is empty and used as a reference.

the individual parameters, namely pH, ionic strength (IS), and the presence of pepsin, a protein-digestive enzyme. As the human GI tract has dynamic media compositions, the INFOGEST protocol was used as a reference [7]: a pH of 3, IS of 98 mM, and presence of pepsin simulates gastric conditions, while a pH of 7 and an IS of 142 mM simulates intestinal conditions. The other combinations were used to discern what is the most impactful parameter and to provide a variety of different media to train the CV; see Table 1. The pepsin concentration was scaled down proportionally using the same INFOGEST method as a reference. For all recordings with pepsin a pepsin concentration of 310 U/ml was employed.

Table 1: Media composition table for swelling tests, protein content loss, and video recordings and analysis. I stands for "intestinal" ionic strength (142 mM); S stands for "stomach" or "gastric" ionic strength (98 mM). The ionic strength was achieved through NaCl (and a small amount of CaCl₂ for media with pepsin) dissolution.

Media Acronym	pH	Ionic Strength (IS)	Pepsin Presence
70	7	0	-
7S	7	S	-
7I	7	I	-
30	3	0	-
3S	3	S	-
3I	3	I	-
3Sp	3	S	yes
1.5Sp	1.5	S	yes

2.5 Swelling Test

Gel cylinders (1.2 cm in diameter, 0.95 cm in height) were immersed in beakers in a 37°C bath containing 20 ml of different media compositions (see Table 1) for 0, 15, 30, 60, and 180 minutes. After the given period, the swollen cylinders were retrieved by sieving the solution, carefully blotted dry, and subsequently weighed. The weight of the swollen cylinders was recorded, after which they were dried in an oven incubator overnight to determine the remaining dry weight. Each sample was replicated 3 times, except for those containing pepsin at 180 minutes. No stirring was applied to the media.

The water uptake was measured using the formula from Haro-Gonzalez, *et al.* [25]:

$$Q_s = \frac{W_s - W_d}{W_d}$$

Where W_s and W_d are respectively the swollen and dry masses of the pellets. This same formula was applied for gels in a pepsin media too, because, even though it does not take into account the mass loss and the consequent loss in absolute water retention capacity, it is still a good approximation of the total water retention by dry mass of the sample.

The swelling test was also employed to calculate the expansion of the gel on the gels in front of the microscope camera, later employed for recordings as:

$$\Delta V_{\text{swelled test gel}} / V_{\text{test gel}} = \Delta h_{\text{camera gel}} \cdot A_{\text{camera gel}} / V_{\text{camera gel}}$$

2.6 UV/Vis and Protein Loss Determination

UV/Vis spectroscopy was employed to know how much gelatin was digested and consequently dissolved in the media during the swelling analysis in the presence of pepsin (1.5Sp and 3Sp). To do so, media was collected at each time point during the swelling test (0, 15, 30, 60, and 180 minutes), pepsin was deactivated by heat, and the solution was analyzed in the spectrometer.

Considering that both the yellow dye and the gelatin absorbed between 200 and 300 nm, a wavelength of 235 nm was used to read absorbance, as it was the best trade-off to stay in the linear regime of the Lambert-Beer law and having the smallest ratio between the dye and gelatin absorbance. This will inevitably bring an intrinsic overestimation of the gelatin content due to the yellow dye dissolving in the solution at the same time.

The amount of gelatin digested in the solution during the swelling tests was then used to calculate the theoretical mass loss in the same media of the gelatin gel in front of the microscope camera, proportionally to the volumes in Table 3.

2.7 Video Recordings

Either one or two USB camera microscopes (Figure 3) embedded with gel were put in 250 ml of media with different compositions (Table 1). All the recordings were carried out at the physiological temperature of 37°C in the neutral booth setup seen in Figure 5. A white sheet of paper was used to keep the background constant and to reduce reflection, an ambient light source was used to normalize the lighting of the recordings (Streamplify Light 10 ring light) and the USB camera microscopes were set to a 90° angle with respect to the beaker. The USB camera microscopes were connected to a computer and each video was recorded for 10 hours on OBS (Open Broadcaster Software).

For the video recording with the camera pill, only one recording was carried out. The camera pill cast in gel (see Figure 4) was put in a beaker with 3Sp-media. The beaker was covered in a white sheet of paper to look more similar to the USB camera microscope recordings. The recording was carried out for 4 hours.

2.8 Image Analysis

In order for the 10-hour recordings to be manageable to do any data analysis on, the recordings were reduced in size. This was done through

```
ffmpeg -an -i "$1" -vf framestep=200,setpts=N/30/TB -r 30 "$1".timelapse.mp4
```

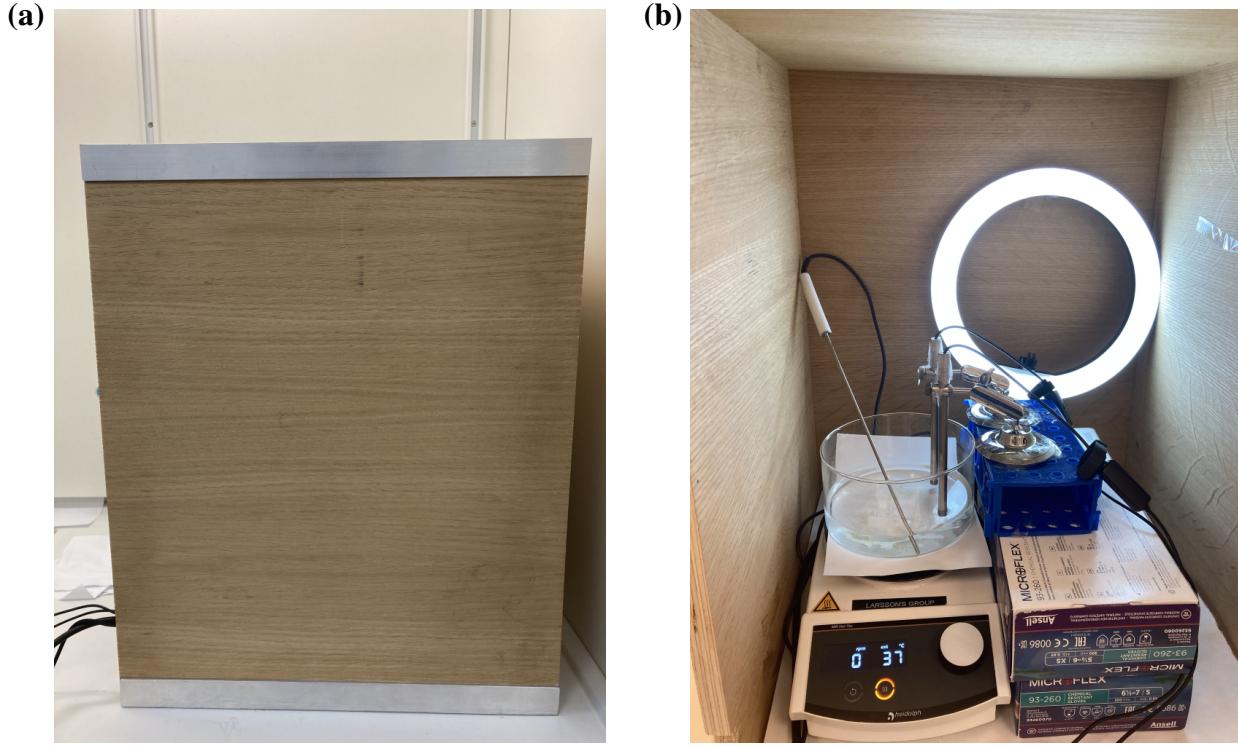


Figure 5: The neutral booth setup in which recordings were performed, with (a) and without lid (b). The heating plate was used to keep the media at 37°C and the white paper sheet to keep the recording’s background constant; the light on the back was used to have constant lighting in the videos.

where only 1 of every 200th frame is kept to create a timelapse video. The chemical process is slow enough so that one frame every few seconds is sufficient as there are no quick changes. This reduces the compute load for further steps.

The subsequent image analysis starts by downscaling to achieve acceptable performance, while cropping out the less important borders of the image. It then reduces irrelevant signal in the image by identifying and masking out air bubbles, before calculating average HSV saturation as a time series. Finally, and as I will motivate, unfortunately, we are forced to filter out discontinuities before plotting the time series complemented with an overlaid linear regression. Detailed description and motivation of these steps in the analysis appear in their respective sections within the results and discussion section.

3 Results and Discussion

The following section presents the results and discusses their implications. Findings are drawn from both the gel characterization part of the project and the image analysis, which quantifies gel dissolution and color changes over time. The section concludes with an evaluation of the gel digestion process and the *in vitro* protocol applied in this project.

3.1 Gel Characterization

To characterize the properties of the gelatin-based gels, compression tests, swelling assessments, and UV/Vis spectroscopy were carried out to assess their structural integrity, swelling behavior, and degradation under

physiological conditions. These analyses provide insights into the stability and resilience of the gels, particularly examining TGase's crosslinking function and the degradation effects of pepsin. This characterization lays the groundwork for the primary computer vision study of gel digestion.

3.1.1 Compression Test

Compression tests were carried out to further characterize the gels' strength and evaluate the TGase enzyme's correct action. As observable in Table 2, it is clear that the TGase has effectively crosslinked the gelatin physical gel into a stronger covalently bonded gel which can resist melting at the body temperature of 37°C. Depending on the source, gelatin has a melting point between 30 and 35°C. We, therefore, needed the crosslink to make the gel resistant to the experiment temperature.

Table 2: Strain at break and compression stress at break for gelatin gels with and without the presence of TGase as crosslinking agent.

Gel	Strain at break (mm/mm)	Compression stress at break (kPa)
Gelatin	0.56 ±0.05	9.1 ±3.5
Gelatin + TGase	0.80 ±0.00	23.4 ±6.6

3.1.2 Swelling Tests

Even though infinite swelling is prevented by the network crosslinking, the water uptake potential of gelatin is big despite the gel being made up almost completely of water already [18]. The stomach juice is acidic and in an acidic medium, gelatin type A will be positively charged [26]. The expansion of hydrogels is mainly driven by polymer–water interactions, electrostatic forces, and osmosis and we expect that in cationic hydrogels, thus the degree of swelling will strongly depend on the pH of the surrounding medium and its ionic strength (IS) [18].

Comparing media with the same pH but different IS in Figure 6, the Qs vary greatly within the graphs, for both pH 3 and 7, swelling more the lower the IS is, as expected. In Figure 7, instead, the three different IS are compared for the two different pHs: the differences between the two pHs are not remarkable, except for IS 0 (Figure 7 a), where pH is the only driving force for the swelling and the osmotic pressure is greater than in the other cases; here pH 3 has a greater swelling than pH 7.

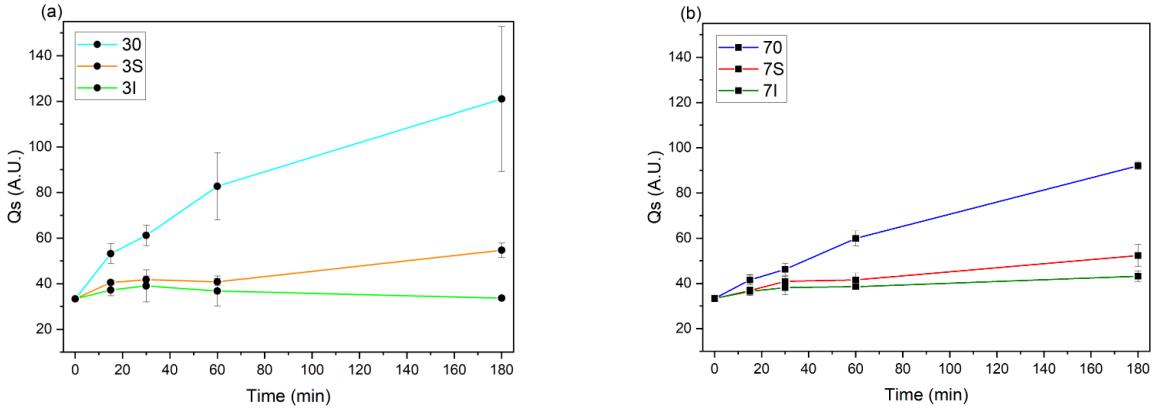


Figure 6: Qs of three different IS (0, S, and I) at two different pH: pH 3 (a) and pH 7 (b).

It is worth mentioning that the presence of pepsin does not influence the absolute value of swelling per dry matter in the gel, plotted in Figure 8. In Figure 8 the media without pepsin at pH 3 and IS S (98 mM) is compared with 3Sp as well as a more acidic media with pepsin (1.5Sp). Despite the fact that Q_s does not take into account the mass loss of gelatin due to pepsin in solution, the absolute water uptake per dry mass (which Q_s represents) remains unchanged.

Furthermore, as introduced in the methods, it is possible to calculate the gel expansion on the microscope probe (Δh) by the test cylinder water uptake in the same media. In Figure 9 we observe a realistic approximation for the 7I case; even though the gel on the camera probe has only one base face in contact with the solvent, while the test cylinder has its whole surface area, the water diffusion is likely fast enough to not matter.

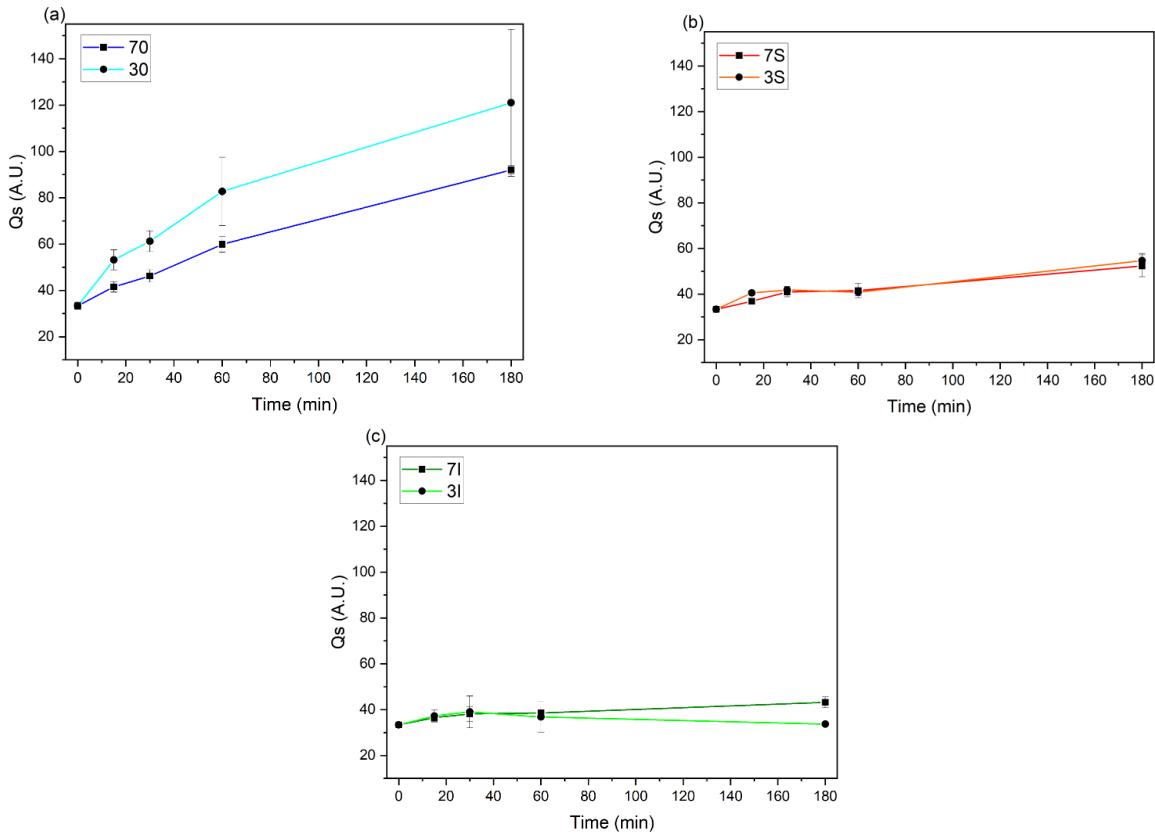


Figure 7: Q_s of the two different pH (3 and 7) at the three different IS (0 (a), S (b), and I (c)).

3.1.3 UV/Vis Spectroscopy

UV/Vis was employed to know how much gelatin was digested and consequently dissolved in the media during the swelling analysis in the presence of pepsin (1.5Sp and 3Sp). This was done also to correlate the mass loss during the recordings with the USB camera microscopes. In Table 3 you can read the volumes of the gels cast in each test cylinder and on the microscope camera tip (which is considerably smaller) and their corresponding content of gelatin, which is also the maximum amount of gelatin that would get digested and leaked in the media during digestion.

The absorbance peak at 235 nm was chosen. However, due to the absorption of both gelatin and the yellow dye at this wavelength, an intrinsic overestimation of approximately ± 6.3 mg for the swelling cylinders

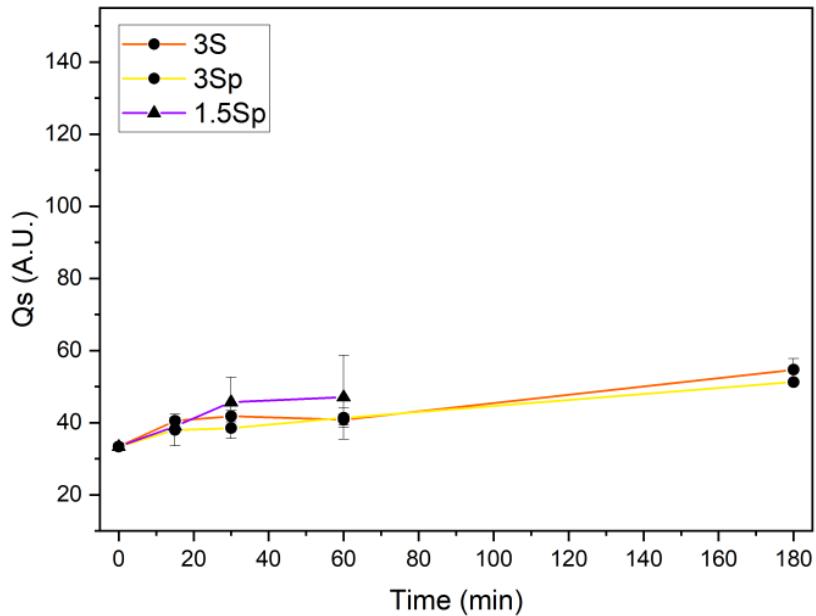


Figure 8: Pepsin influence on pH 3 and 1.5, with or without the presence of pepsin. 1.5Sp was completely digested by 180 min, while 3Sp (180 min) has only one repetition and no standard deviation.

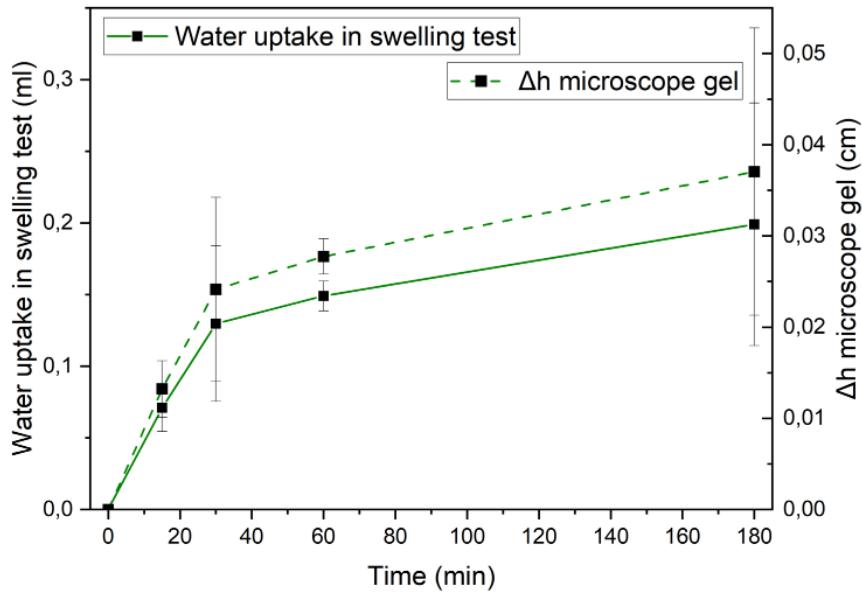


Figure 9: Swelling of 7I plotted with the proportionally calculated height change due to swelling of the probe microscope gel during its first 3 hours of recording in the same media.

and ± 0.45 mg for the microscope camera is expected. A similar release curve to that of gelatin is expected for the yellow dye, leading to a progressively increasing associated error, with the previously stated values representing the maximum possible overestimation. Keeping this in mind in Figure 10, we can observe that the same media composition, but at a lower pH has a faster digestion, therefore a steeper increase in the

graph. This is probably because pepsin's optimal working pH range is between 1.5 and 2.

When translated to the microscope camera, we expect the gelatin to disappear from the video in around 3 minutes for 1.5Sp and 10 minutes for 3Sp, which is untrue. In reality, the gel takes between 1 and 2 hours to be completely digested from the microscope camera tip, thus this approximation does not apply so well as for the water uptake case and overestimates the digestion speed of the gelatin.

Table 3: Table showing the volumes of the gels cast in each test cylinder and on the microscope camera tip and their corresponding content of gelatin, which is also the maximum amount of gelatin that would get digested and leaked in the media during digestion

Casting mold	V of gel cast (ml)	Gelatin present in the gel (mg)
Test cylinders	1.07	33.33
Microscope camera	0.08	2.39

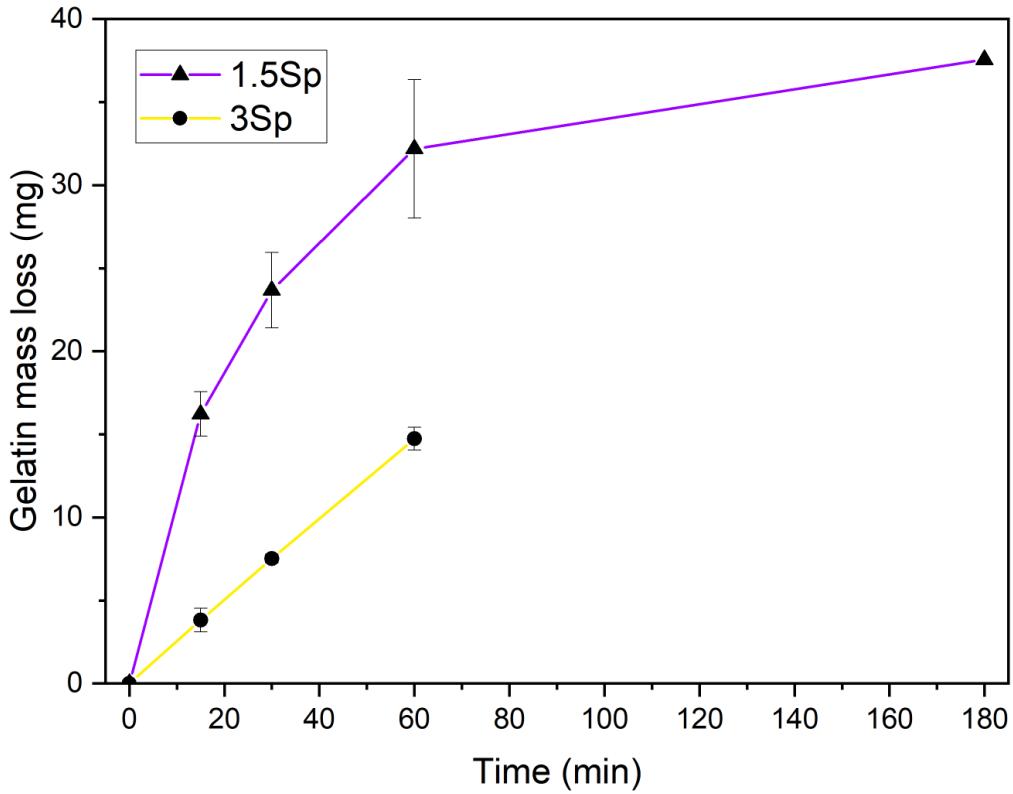


Figure 10: Mass loss of the gelatin cylinder during pepsin digestion calculated by UV/Vis spectroscopy. 1.5Sp has one only repetition at 180 min, while 3Sp has none.

3.2 Steps of Image Analysis

This section explains the steps of image analysis from timelapse to time series, including motivating them based on examples from the data. The entire code in addition to the timelapse data and plots themselves can be found in the GitHub repository at <https://github.com/Juglone/KEM929>.

3.2.1 Clip to Middle Square

Given the large volume of video data, downsampling is essential. We previously discussed temporal downsampling by creating a timelapse in Section 2.8, but it is also necessary to spatially downsample. This is done mainly by downscaling to reduce the resolution, but it can be combined with cropping to remove more pixels for the same reduction in resolution. To optimize data quality, we crop each frame to the central region where visual information is less impacted by dark occlusions along the edges (see Figure 11). This center-crop square, with a side length half the frame's height, allows us to retain critical signal in regards to for example bubble edge resolution while minimizing less relevant data at the periphery.

3.2.2 Brightness vs. Saturation

To effectively track changes in images over time, relevant metrics must be recorded. One approach involves examining the brightness of each frame. However, as shown in Figure 11, the difference between frames 11a and 11b is not solely a shift in absolute brightness but also a variation in saturation or "yellowness" over time.

Observing Figure 11 and the corresponding video, a progressive decrease in yellowness is apparent, which appears to be a consistent trend across recordings. Therefore, quantifying this "yellowness" may provide a more accurate assessment of visual changes than grayscale brightness alone. By analyzing the frames in the HSV color space, it is evident that the saturation value decreases from 124.1 to 98.8, while the brightness shifts from 147.3 to 142.2 (see Table 4 for detailed values). Thus, analyzing the frames in the HSV color space offers a more informative measure of change than brightness alone.

Table 4: Comparison in different color spaces of Frame 1 and Frame 2. The values are given on a scale of 0-255

Frame 1	Frame 2
Mean Brightness: 147.3	Mean Brightness: 142.2
Mean RGB: [164.8, 150.1, 86.8]	Mean RGB: [155.7, 144.4, 96.1]
Mean HSV: [24.5, 124.1, 164.8]	Mean HSV: [24.8, 98.8, 155.7]
Mean HSL: [24.5, 125.9, 98.7]	Mean HSL: [24.8, 126.1, 73.2]



(a) This is a frame from the start of the 10-hour-long recording of the microscope with gel cast in a media with pH 3 and ionic strength S. The recording was carried out on 2024-08-22 (a)



(b) This is a frame from the end of the 10-hour-long recording of the microscope with gel cast in a media with pH 3 and ionic strength S. The recording was carried out on 2024-08-22 (a)

Figure 11: Two frames from the start and end of a recording (3S, 2024-08-22 (a))

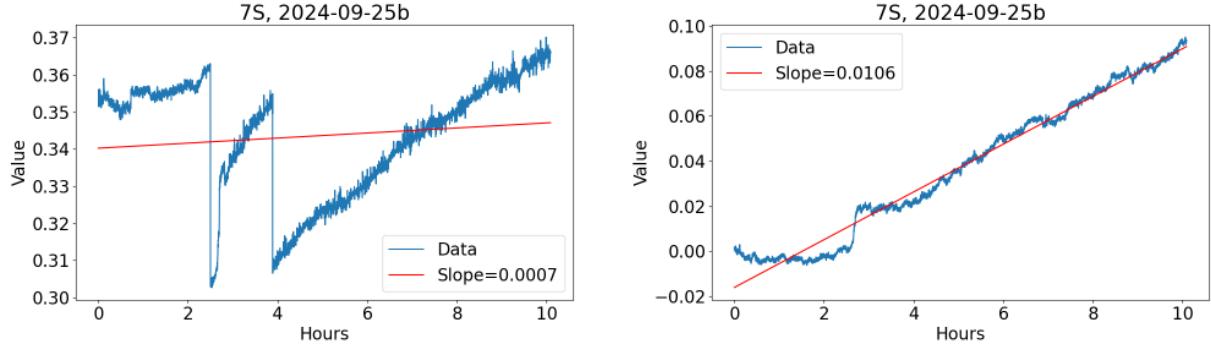
3.2.3 Bubble Removal Using Hough Circle Detection

One might expect the gel to appear as a uniform yellow color on video, but bubbles often form on the interface between the gel and the media. These bubbles move around during the timelapse, diffracting light in a manner seemingly uncorrelated to the decrease in yellowness as the gel dissolves. It is therefore motivated to filter these bubbles out of the video frames before averaging spatially.

This is done using Hough circle detection which is explained in section 1.6, see Figure 1 for an example. To make sure the detected circles fully cover all bubbles, their radius is expanded by 2 pixels before writing them to a mask. Hough circles have significant false positives for noisy images, so we keep the masks for the 40 most recent frames and compute their 20th percentile per pixel in order to reduce their sensitivity. The averaging of yellowness/HSV saturation values is then done ignoring pixels covered by detected bubbles.

3.2.4 Discontinuity Removal

We do not expect the video to include sudden drops or sudden peaks between two neighboring frames. The setup does not suggest any physical explanations as to why sudden drops, such as in Figure 12a, would appear. Looking at the video, a probable solution seems to be auto brightness adjustments in the camera. In order to adjust for this, sudden large changes were clamped with previous values in a given interval. The result from applying this discontinuity removal can be seen in Figure 12.



(a) Graph before implementing discontinuity removal. The graph is for the recording with media with pH 7 and ionic strength S. The experiment was carried out 2024-09-25 (b)

(b) Graph after implementing discontinuity removal. The graph is for the recording with media with pH 7 and ionic strength S. The experiment was carried out 2024-09-25 (b)

Figure 12: Discontinuity removal

3.2.5 Linear Regression

We expect gel dissolution to change the color over time, with a faster change for stronger media configurations. This motivates performing curve fitting to quantitatively measure the strength of the media. It is not obvious whether this relationship would be linear or something else, but we opt for a linear regression since it minimizes overfitting and all relationships are at least locally linear. Additionally, a constant offset is applied so that every graph starts at zero on the y-axis to more clearly highlight the size of the change.

3.3 Gel Digestion and Protocol Evaluation

As we can see in Figures 13-22 there are a lot of recordings. Table 5 shows an overview with a short observation on each trial. In short, the results are all over the place with some plots increasing and some decreasing. We would expect the plots to develop in the same direction over all of the recordings, as the yellow gel is partially dissolved or digested in the different videos. The ambient lighting is also the same for all recordings, pictured in Figure 5. Yet despite this, the plots are wildly different from each other.

Looking at Figure 14 as an example, we can see that the two graphs (a) and (b) from the first 3S-recording (2024-08-26) develop in the same direction, a large linear increase in the metric value. However, the two recordings carried out on the other day of 3S-recording are very different from the first two and even each other. To emphasize, these were recording concurrently immersed in the same beaker and should differ only in things like minor gel imperfections, or as I will go on to argue, undesired camera software processing.

A very significant takeaway from these results however is that the pepsin trials in Figure 19-20 do have sensible results. They show a sharp decrease corresponding to the gel fully dissolving within the first 2 hours. While they diverge after that, the initial trend is clear enough that we can identify the dissolution time clearly. Figure 21 shows frames at the start, after dissolution and at the very end of one of these recordings, showing that the plots indeed capture what is going on in the video accurately. It appears the trends for pepsin trials are clear enough to surface over the admittedly very prominent irrelevant noise.

Table 5: Overview of all gel dissolution plots. Trends are not consistent, with some increasing, some decreasing, some having difficult-to-explain nonlinear jumps. Still, almost all trials experience significantly larger changes than the control with no gel.

Figure	Media	Magnitude	Observations on trials
13	3O	0.05-0.15	4x non-monotonic changes*
14	3S	0.06-0.25	2x linearly increasing, 2x nonlinearly decreasing*
15	3I	0.025-0.4	3x stationary, linearly decreasing. Change in slope at 8 hours could indicate full dissolution, but in the video gel is in fact still very visible.
16	7O	0.05-0.25	3x stationary, linearly decreasing.
17	7S	0.05-0.2	stationary, 2x linear decrease, nonlinear increase*.
18	7I	0.03-0.15	2x linear increase, stationary, linear decrease
19	1.5Sp	0.1	Clear linear decrease within first 2 hours as the gel dissolves fully.
20	3Sp	0.15	Clear linear decrease within first 2 hours as the gel dissolves fully.
22	7O additional	0.1	nonlinear decrease
23	No gel	0.015-0.03	2x linear increase

I claim that the issues with data quality can largely be explained by the non-reproducibility of the camera. One piece of evidence of this is the significance of the discontinuity-removal in the data analysis (see Section 3.2.4) – without it, many measurements are dominated by unpredictable large jumps. From also looking at these videos in detail, I am convinced that these changes are due to software post-processing within the camera, as they affect the frame globally in a sudden way.

In particular, the camera microscope used is not color-accurate, meaning that the same physical scene does not always give the same digital frame. Despite attempting to disable auto-brightness and auto-exposure in OBS, similar post-processing is evidently present and makes the output videos much more difficult to work with from a computer vision perspective.

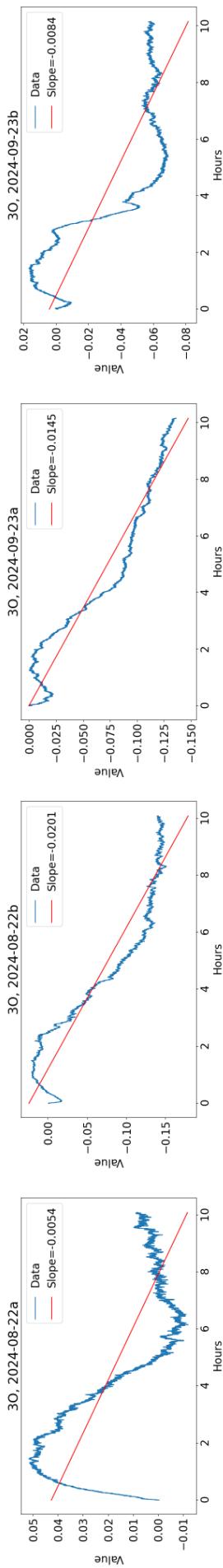


Figure 13: Time series of filtered average saturation for 3O trials

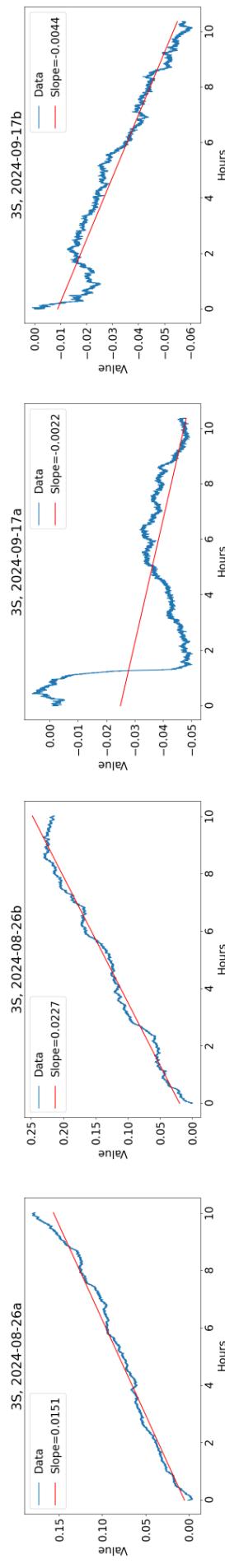


Figure 14: Time series of filtered average saturation for 3S trials

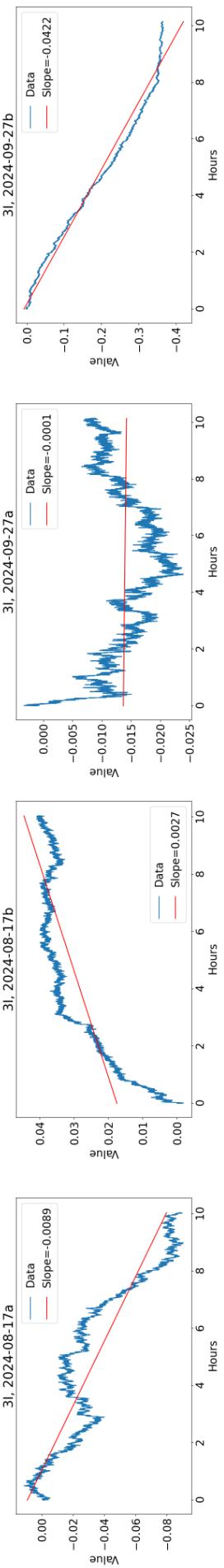


Figure 15: Time series of filtered average saturation for 3I trials

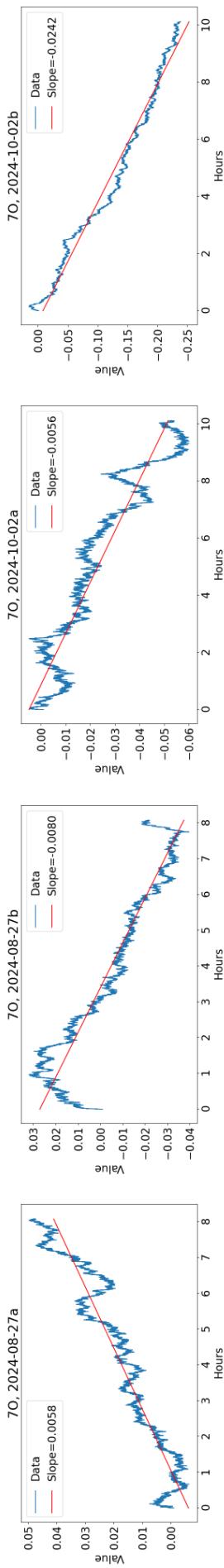


Figure 16: Time series of filtered average saturation for 7O trials

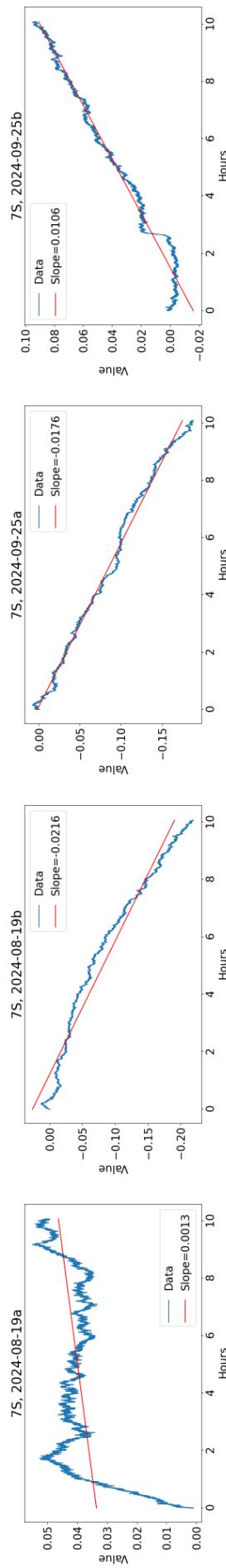


Figure 17: Time series of filtered average saturation for 7S trials

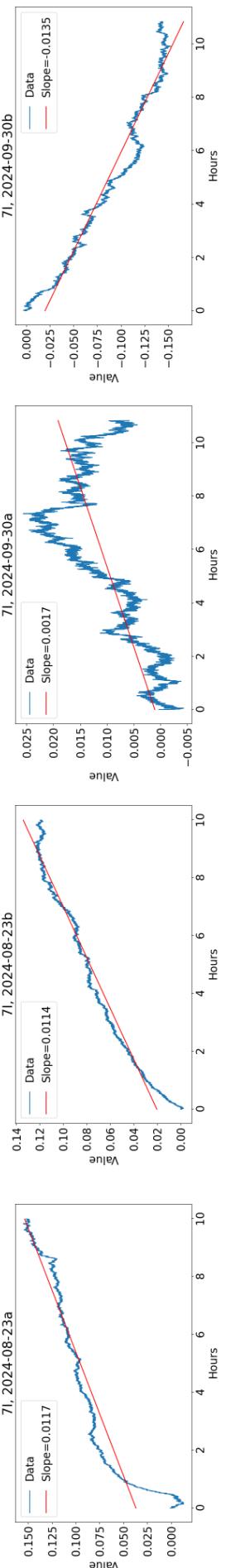


Figure 18: Time series of filtered average saturation for 7I trials

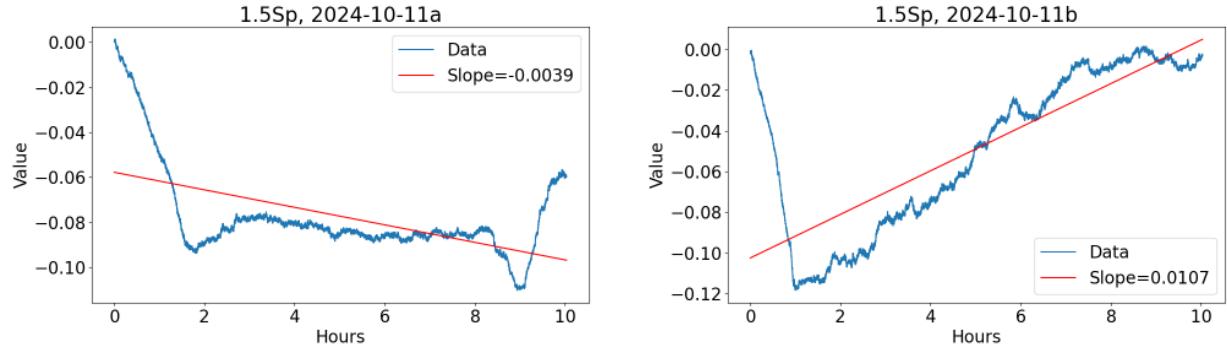


Figure 19: Time series of filtered average saturation for 1.5Sp trials

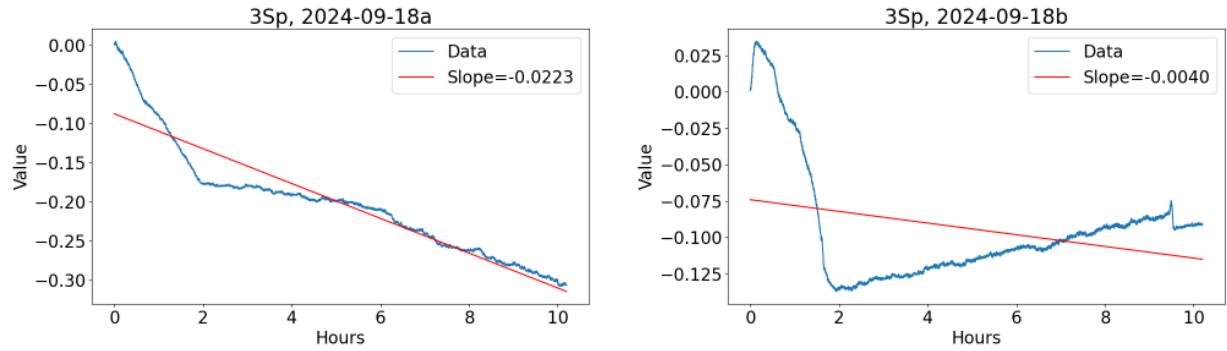


Figure 20: Time series of filtered average saturation for 3Sp trials



Figure 21: Recording 3Sp, 2024-09-18a, first frame, after two hours and last frame. The clear yellow tint disappears as the gel dissolves, with the camera focusing directly on the sheet of paper placed under the beaker. Out-of-focus objects or artifacts are visible along the borders of both frames.

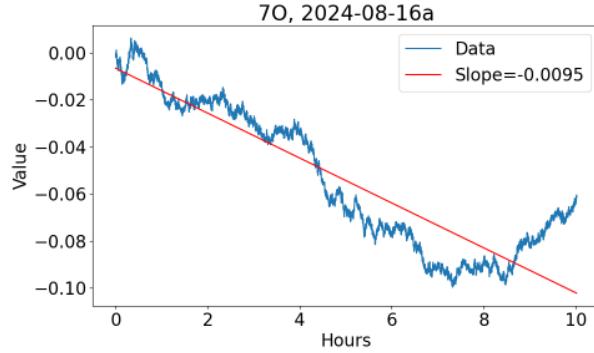


Figure 22: Time series of filtered average saturation of an additional 7O recording

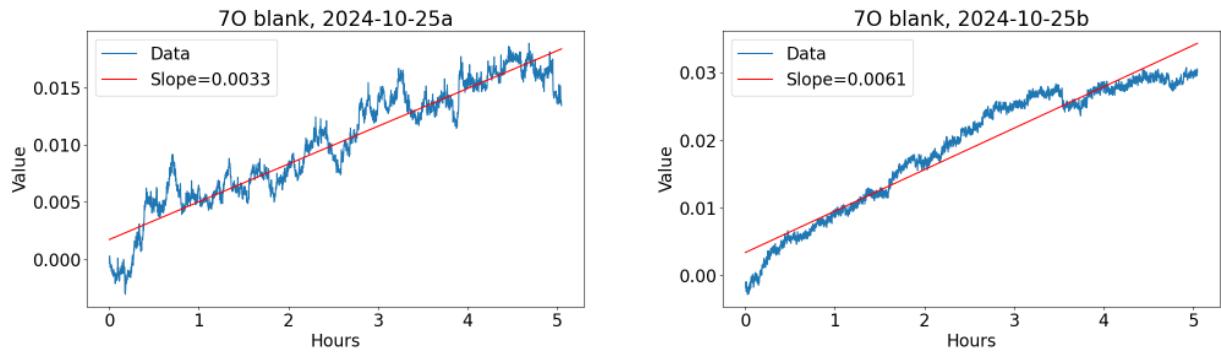


Figure 23: How recordings change over time when there is no gel present. The magnitude of this change is significantly smaller than the change in most other recordings, indicating that the difficult-to-explain movements in other recordings is not plain ambient noise.

Figure 24 displays a frame from the camera pill recording, where internally lit camera pills illuminate the surroundings. The frame differs significantly from those captured with the USB camera microscope under ambient light. The primary factor driving this difference is the different lighting conditions, internal versus ambient light, which complicates direct comparison between frames from the two setups. However, the use of an internally lit camera pill offers other opportunities to the experimental setup, such as measuring reflection and imprinting gels with pattern – both of which are explained in more detail in Section 4.



Figure 24: A frame from the camera pill recording. It records a full 360° view, which due to its gel casting frame appears as 4 sections. Sections 1, 2 and 4 are covered in gel while the third is not have serves as a reference.

4 Conclusion

This thesis set out to develop a proof of concept for quantitatively assessing the digestion process using cameras, including developing an *in vitro* protocol and computer vision code, and to evaluate whether this concept is feasible and where the challenges lie. This proof of concept was developed but the final results tell a mixed story dominated by the realization that color-accurate work require color-accurate cameras but also raising questions about color-accuracy-resistant alternatives and the sensibility of looking too much into pepsin-less media in the first place. But overall the results do indicate that at least in the case of a presence of pepsin, the point at which the gel is fully dissolved can be identified using this thesis' protocol.

Compression tests are shown to be a reliable analysis when it comes to crosslinking evaluation. Both the swelling tests and the UV/Vis spectroscopy are very useful techniques to evaluate the water uptake and the mass loss of the gels during the immersion in the media. Swelling is particularly useful when it comes to approximating the height change in front of the camera during the recording, and it could be easily integrated to the data processing. UV/Vis spectroscopy is a solid technique to determine how much gelatin has gotten released in the media during digestion, but it appeared to be hardly transposable to the microscope camera case, probably due to the yellow dye interference in the absorbance and the tiny quantity of gel and gelatin mass in comparison with the swelling test cylinders.

As seen in the results, there was no clear correlation between the dissolution estimate provided by the image processing and the time since immersing the gel in media. Additionally there were significant differences between two gels immersed in the same media in the same beaker at the same time. These indicate underlying problems in data quality that the image processing was unable to correct.

Data quality is affected by multiple factors including the reproducibility of the gel, the media, and the experimental setup and data collection. Media, gel casting, and experimental setup seem unlikely to have a significant effect since the protocol employed is very precise. Therefore, it is likely that the data collection is the weaker link of the chain. As previously mentioned, factors such as the necessity of discontinuity removal are very strong evidence towards the camera and its lack of color accuracy and the impossibility of fully disabling on-device software processing of the image, including auto-exposure, as the underlying problem with the data.

There are ways of changing the experimental setup and image processing to be robust against changes in color balance and exposure. One way is to simply collect significantly more data, and use deep learning to pick up on more subtle clues in the videos. But this approach makes it possible to introduce significant biases making the model unusable (for example, imagine a correlation between day-of-week and gel age and media configuration), and more importantly, collecting significantly more data would be an unreasonable logistical hurdle for as long as a single measurement occupies one camera for one night.

Another way of increasing robustness to camera color inaccuracy is to instead consider patterns and contrast, which are significantly less affected by color balance. If one could imprint gels with a pattern, for example, a chessboard pattern, dissolution could be correlated with the smearing of the pattern. This approach seems to me very promising but has hurdles in how the gel would be manufactured. Perhaps, 3D printing of hydrogels could be employed.

On a closing note, I would like to speculate on opportunities offered by using internally lit camera pills. While measuring the tint of the colored gel and correlating with its dissolution should remain possible, some new techniques would be unlocked in addition. A pill consisting of a camera and light, enclosed in transparent plastic, covered in fairly opaque gel material, would mostly pick up light reflected in the plastic on its way out. As the gel dissolves, this would transition to most light being reflected in the pill's surroundings. Hence a computer vision algorithm could look for the fading of its own reflection as it is replaced by diffuse light from its surrounding. This is potentially easier than measuring gel-induced tinting of ambient light and would not require color accuracy.

References

- [1] Y. Wang, Y. Huang, R. C. Chase, *et al.*, “Global burden of digestive diseases: A systematic analysis of the global burden of diseases study, 1990 to 2019,” *Gastroenterology*, vol. 165, no. 3, pp. 773–783, 2023.
- [2] J. M. Ballreich, C. P. Gross, N. R. Powe, and G. F. Anderson, “Allocation of national institutes of health funding by disease category in 2008 and 2019,” *JAMA Network Open*, vol. 4, no. 1, e2034890–e2034890, 2021.
- [3] R. Lucas-González, M. Viuda-Martos, J. A. Pérez-Alvarez, and J. Fernández-López, “In vitro digestion models suitable for foods: Opportunities for new fields of application and challenges,” *Food Research International*, vol. 107, pp. 423–436, 2018, ISSN: 0963-9969. DOI: <https://doi.org/10.1016/j.foodres.2018.02.055>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0963996918301480>.
- [4] I. Sensoy, “A review on the food digestion in the digestive tract and the used in vitro models,” *Current Research in Food Science*, vol. 4, pp. 308–319, 2021, ISSN: 2665-9271. DOI: <https://doi.org/10.1016/j.crfs.2021.04.004>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2665927121000307>.
- [5] K. Verhoeckx, P. Cotter, and I. López-Expósito, *Dynamic Digestion Models: General Introduction*. Cham (CH): Springer, 2015.
- [6] O. Antal, I. Dalmadi, and K. Takács, “Upgrading in vitro digestion protocols with absorption models,” *Applied Sciences*, vol. 14, no. 18, p. 8320, 2024.
- [7] A. Brodkorb, L. Egger, M. Alminger, *et al.*, “Infogest static in vitro simulation of gastrointestinal food digestion,” *Nature protocols*, vol. 14, no. 4, pp. 991–1014, 2019.
- [8] K. Dawson-Howe, *A practical introduction to computer vision with opencv*. John Wiley & Sons, 2014.
- [9] A. Gupta, “Current research opportunities for image processing and computer vision,” *Computer Science*, vol. 20, pp. 387–410, 2019.
- [10] X. Su, X. Yan, and C.-L. Tsai, “Linear regression,” *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 4, no. 3, pp. 275–294, 2012.
- [11] J. Canny, “A computational approach to edge detection,” *IEEE Transactions on pattern analysis and machine intelligence*, no. 6, pp. 679–698, 1986.
- [12] G. Bradski, “The OpenCV Library,” *Dr. Dobb’s Journal of Software Tools*, 2000.
- [13] J. Illingworth and J. Kittler, “The adaptive hough transform,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, no. 5, pp. 690–698, 1987.
- [14] S. Süsstrunk, R. Buckley, and S. Swen, “Standard rgb color spaces,” in *Color and imaging conference*, Society of Imaging Science and Technology, vol. 7, 1999, pp. 127–134.
- [15] C. A. Poynton, “Rehabilitation of gamma,” in *Human Vision and Electronic Imaging III*, SPIE, vol. 3299, 1998, pp. 232–249.
- [16] S. Romani, P. Sobrevilla, and E. Montseny, “Variability estimation of hue and saturation components in the hsv space,” *Color Research & Application*, vol. 37, no. 4, pp. 261–271, 2012.
- [17] M. Djabourov, “Gels,” 2020.
- [18] K. Omidian Hossein; Park, “Introduction to hydrogels,” 2010.

- [19] P. Navard, *The European Polysaccharide Network of Excellence (EPNOE): research initiatives and results*. Springer Science & Business Media, 2012.
- [20] S. Mohanto, S. Narayana, K. P. Merai, *et al.*, “Advancements in gelatin-based hydrogel systems for biomedical applications: A state-of-the-art review,” *International Journal of Biological Macromolecules*, p. 127 143, 2023.
- [21] Ö. Pekcan and S. Kara, “Gelation mechanisms,” *Modern Physics Letters B*, vol. 26, no. 27, p. 1 230 019, 2012.
- [22] A. Totosaus, J. G. Montejano, J. A. Salazar, and I. Guerrero, “A review of physical and chemical protein-gel induction,” *International journal of food science & technology*, vol. 37, no. 6, pp. 589–601, 2002.
- [23] N. Bosnjak, S. Nadimpalli, D. Okumura, and S. A. Chester, “Experiments and modeling of the viscoelastic behavior of polymeric gels,” *Journal of the Mechanics and Physics of Solids*, vol. 137, p. 103 829, 2020.
- [24] V. K. Thakur and M. K. Thakur, *Polymer Gels*. Springer, 2018.
- [25] J. N. Haro-González, B. N. S. de Alba, N. Morales-Hernández, and H. Espinosa-Andrews, “Type a gelatin-amidated low methoxyl pectin complex coacervates for probiotics protection: Formation, characterization, and viability,” *Food Chemistry*, vol. 453, p. 139 644, 2024, ISSN: 0308-8146. DOI: <https://doi.org/10.1016/j.foodchem.2024.139644>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0308814624012949>.
- [26] C.-L. Tseng, K.-H. Chen, W.-Y. Su, Y.-H. Lee, C.-C. Wu, and F.-H. Lin, “Cationic gelatin nanoparticles for drug delivery to the ocular surface: In vitro and in vivo evaluation,” *Journal of Nanomaterials*, vol. 2013, no. 1, p. 238 351, 2013.