



télécom
saint-étienne

Telecom Saint-Étienne
FISE 2



MALLOWMETER

– Données –

ÉTUDE EN FILIÈRE INGÉNIEUR SOUS STATUT ÉTUDIANT
11.02.2025

MAXIME JOURNOUD LUCAS LESCURE AUBIN SIONVILLE

Table des Matières

I. Rappel du projet	3
II. État de l'art	3
III. Définition de la base de données	3
IV. Procédures d'acquisition	4
V. Critères d'évaluation	4
VI. Métriques d'évaluation	5

I. Rappel du projet

L'objectif est de classer des photos de marshmallows blancs en fonction de leur niveau de cuisson. L'application devra fournir, à partir d'une photo de marshmallow et d'un masque binaire représentant la position du marshmallow dans l'image, une étiquette correspondant à sa classe parmi les 4 degrés de cuisson possibles.

II. État de l'art

N'ayant pas trouvé de base d'images de marshmallows répondant à nos besoins (différents degrés de cuisson, nombre important d'images, diversité des cuissons), nous avons décidé de la constituer nous-mêmes, principalement en réunissant des clichés trouvés sur le web. En parallèle, la recherche d'images sur Internet sera complétée par une acquisition manuelle afin de garantir un nombre conséquent d'échantillons et une diversité dans les cuissons.

À ce jour, nous disposons de **68 images de marshmallows acquis depuis le web** et **0 par acquisition manuelle**, réparties selon le tableau ci-dessous.

Degré de cuisson	Images Web	Images Manuelles
Pas cuit	13	0
Peu cuit	6	0
Bien cuit (sur-représenté)	37	0
Trop cuit	12	0

Table 1. Répartition des images selon le degré de cuisson

Afin d'équilibrer notre base, nous prévoyons de réaliser **au minimum 29 acquisitions manuelles d'images supplémentaires** de marshmallows cuits à différents niveaux afin d'amener chaque classe au compte de 20 unités, et nous réaliserons un tri parmi les images de classe sur-représentée.

On aura éventuellement recours à l'enrichissement dans le cas où les données seraient insuffisantes, notamment dans un second temps pour la classification CNN.

III. Définition de la base de données

Notre base de données sera composée de **4 classes**, avec **20 images minimum par classe**, soit un total d'au moins **80 images**. Chaque image sera accompagnée de son **masque binaire**, indiquant la position du marshmallow sur l'image.

Les images seront toutes placées dans un dossier commun et respecteront les critères suivants :

- **Format** : PNG 512×512 pixels
- **Fond** : Neutre, avec le marshmallow au premier plan
- **Type de marshmallow** : Principalement de couleur blanche

Les fichiers seront nommés selon la structure suivante, où [X] est le numéro de l'échantillon et [1-2-3-4] la classe déterminée manuellement :

- **Images** : `img[X]_i_[1-2-3-4].png`,
- **Masques** : `img[X]_m.png`, correspondant au masque associé à l'image.

IV. Procédures d'acquisition

Les images collectées sur internet seront choisies arbitrairement en prenant compte du détachement des marshmallows par rapport à leur fond, et par rapport à la cuisson. On évitera de prendre des images trop pixelisées et on fera attention à vérifier les droits d'images afin d'être sûrs qu'elles soient libres d'utilisation.

D'autre part, les images qui seront prises à la main seront prises en faisant attention à ce que le fond puisse également bien se détacher.

Nous effectuerons ensuite une **suppression de l'arrière-plan** de façon à n'avoir que le marshmallow comme objet d'intérêt. Ceci sera réalisé à l'aide d'outils comme RemoveBG ou RemovalAI. Grâce à cette étape, nous pourrons construire un **masque du marshmallow**, qui sera une donnée supplémentaire à utiliser, notamment pour la classification via CNN. Les images seront ensuite **rognées**, de façon à obtenir un ratio objet/image d'environ 30 à 75% sur toutes les images. Enfin, on réalisera de façon adéquate une **mise à l'échelle** pour s'assurer que la taille des images soit toutes au format 512x512 pixels.

V. Critères d'évaluation

Nous allons tout d'abord effectuer une **analyse descriptive des données** afin d'évaluer la distribution des caractéristiques visuelles des images et détecter d'éventuels déséquilibres dans notre base. Cette analyse portera sur les aspects suivants :

- **Moyenne et écart-type des composantes L et b dans l'espace Lab pour chaque classe** : L'espace colorimétrique Lab permet une meilleure représentation des variations de couleur et de luminosité. La composante **L** correspond à la luminosité (du noir au blanc), tandis que la composante **b** représente la balance entre les tons jaunes et bleus, particulièrement pertinente pour identifier les degrés de cuisson des marshmallows. Nous calculerons la **moyenne et l'écart-type** de ces valeurs pour chaque classe afin de vérifier si les classes sont bien différenciées en termes de couleur et si certaines d'entre elles présentent une trop grande **variabilité intra-classe**.
- **Proportion de pixels blancs dans les masques** : Chaque image est accompagnée d'un masque binaire indiquant la position du marshmallow. La proportion de pixels blancs dans le masque permet d'estimer la **taille relative du marshmallow dans l'image**. Nous analyserons la distribution de cette proportion dans chaque classe pour nous assurer que les tailles des marshmallows sont homogènes et qu'il n'existe **pas de biais majeur entre les classes** (ex. marshmallows plus gros pour un certain degré de cuisson).

Ces analyses nous permettront de détecter d'éventuels déséquilibres et de prendre les mesures nécessaires pour améliorer notre base de données. Il est également essentiel d'évaluer la qualité de notre base de données afin de garantir des performances correctes du modèle de classification. Deux critères sont particulièrement importants :

- **La représentativité** : notre base doit être variée, afin de couvrir un **éventail assez large de cas possibles** et d'**éviter un modèle spécifique à un contexte particulier**. Une mauvaise représentativité risquerait de réduire la capacité de généralisation du modèle et d'entraîner une mauvaise classification sur des données inconnues.
- **Équité** : chaque classe doit être **représentée de manière équilibrée** afin d'éviter des biais dans l'apprentissage. Une classe sur-représentée pourrait amener le modèle à privilégier certaines prédictions au détriment des autres, ce qui réduirait sa précision globale.

VI. Métriques d'évaluation

- **Évaluation de l'équilibre des classes** : Notre base de données est composée de $k = 4$ classes, notées $i \in \{1, 2, 3, 4\}$, contenant chacune N_i échantillons, pour une proportion $p_i = \frac{N_i}{\sum_{j=1}^k N_j}$. Notre mesure de l'équilibre entre les classes sera **l'entropie de Shannon normalisée**, définie par :

$$H_n = -\frac{1}{\log_2(k)} \sum_{i=1}^k p_i \log_2(p_i),$$

où H_n varie entre 0 (base très déséquilibrée) et 1 (répartition parfaitement uniforme). Nous considérons que la base est suffisamment équilibrée si $H_n > 0.75$.

- **Vérification de la diversité intra-classe** : Au-delà de la répartition globale des échantillons, il faut s'assurer que **chaque classe couvre une variété suffisante de cas** pour éviter un **sur-apprentissage** sur des caractéristiques trop spécifiques. Pour cela, nous analysons la variation des conditions de prise de vue: les images doivent inclure **différentes directions d'éclairage et angles de vue** afin de rendre le modèle robuste aux variations du contexte. Cette partie sera vérifiée empiriquement au vu de la difficulté de l'établissement d'un critère de mesure objectif.

L'objectif est d'éviter que le modèle ne se base sur des artefacts visuels non pertinents (ex : un fond particulier présent majoritairement dans une classe) et qu'il puisse généraliser efficacement à des images inconnues.