

Reconnaissance des yeux pour la détection de la température du canthus interne par l'intelligence artificielle : YOLOv5

Malak Ghourabi, Farah Mourad-Chehade, Aly Chkeir
Laboratoire Informatique et Société Numérique (LIST3N)
Université de Technologie de Troyes
Troyes, France
{malek.ghourabi, farah.chehade, aly.chkeir}@utt.fr

Abstract - Les infections virales constituent un danger mortel pour les individus âgés et physiquement fragiles. Étant donné l'augmentation substantielle de cette population en Europe, il est crucial de détecter précocement la fragilité physique et les maladies infectieuses. Dans cet article, nous introduisons une méthode de détection de la température corporelle sûre, exacte et rapide, qui pourrait être intégrée dans les foyers ou les résidences pour personnes âgées. Cette méthode vise à identifier l'un des symptômes des maladies infectieuses : une température corporelle élevée. Nous avons concentré nos efforts sur la reconnaissance des yeux dans des images thermiques du visage, puis avons analysé la région des yeux détectés pour déterminer la température du canthus interne. La détection des yeux a été effectuée en entraînant quatre variantes de l'algorithme de détection d'objets You Only Look Once 5ème version (YOLOv5) : nano, petit, moyen et grand. Nous avons utilisé un ensemble de données de 4255 images thermiques pour l'entraînement, combinant deux jeux de données distincts et en appliquant des techniques d'augmentation des données. Les résultats ont révélé une précision moyenne mAP égale à (99,5 %) pour les différents modèles entraînés. Le modèle YOLOv5 grand s'est avéré être le plus rapide, fonctionnant à 115 FPS.

Keywords: maladies infectieuses ; COVID-19 ; fragilité physique ; personnes âgées ; caméra thermique ; YOLO ; reconnaissance d'objets ; détection de la température ; canthus interne.

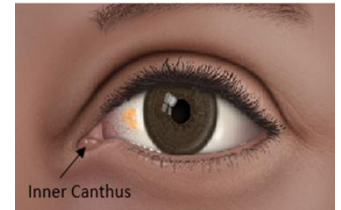
I. INTRODUCTION

Selon la conférence ministérielle sur le vieillissement (UNECE), les personnes âgées représenteront un quart de la population européenne, soit une augmentation de 70 % en 20 ans [1]. La population âgée est plus vulnérable aux infections virales, avec des taux de morbidité et de mortalité plus élevés [2]. Il est crucial de détecter rapidement la fragilité physique et les signes des maladies infectieuses pour intervenir plus tôt et prévenir les conséquences mortelles. La fragilité physique est un syndrome clinique fréquent chez les personnes âgées, augmentant le risque de problèmes de santé. L'échelle de Fried évalue cinq critères pour détecter la fragilité : perte de poids involontaire, vitesse de marche lente, épuisement général, faible activité physique et faible force de préhension. Des

outils de dépistage sont utilisés pour identifier la fragilité physique, comme ceux proposés par le projet ARPEGE. Ce projet offre un ensemble de capteurs installés au domicile des personnes âgées : (1) une balance pour analyser l'équilibre et la perte de poids, (2) une grip ball, qui est une balle avec un capteur de force pour étudier la réduction de la force de préhension maximale, (3) deux questionnaires pour détecter l'épuisement général, (4) un radar pour mesurer et suivre la vitesse de marche, et (5) un questionnaire pour détecter toute baisse de l'activité physique [3]. Parallèlement à la détection de la fragilité, il est essentiel d'identifier le risque d'infection virale pour recevoir un traitement rapide. Dans le cas de COVID-19, par exemple, les personnes les plus touchées étaient celles âgées de 65 ans et plus [4].

Trois indicateurs principaux permettent de suspecter la présence d'une infection virale : (1) la fièvre (élévation de la température), (2) la fatigue et (3) la toux [5]. La caméra thermique infrarouge s'est avérée être un outil fiable et sans contact pour la détection de la fièvre, contribuant également à la prévention de la contamination. Pour mesurer la température corporelle avec une caméra thermique, il est important de se baser sur la température de la région du canthus interne (figure 1). Le canthus interne de l'œil est le coin de l'œil où les paupières se rejoignent, qui s'est avéré être la région optimale pour le dépistage de la température en raison du passage de l'artère carotide interne sous ce canthus [6].

Des travaux de recherche ont été menés pour détecter la température du canthus interne, tels que ceux présentés dans [7], basés sur l'apprentissage de deux modèles de l'algorithme de détection d'objets "You Only Look Once" (YOLO), YOLOv4 et YOLO-tiny, avec 606 images thermiques. Les modèles ont obtenu des scores de précision de 94 % et 99 % et un score de rappel de 99 %. De plus, Budzan et al. ont localisé les yeux pour détecter la température du canthus interne en utilisant la transformée de Hough aléatoire pour les détections d'ellipses. Cette méthode a été testée sur 125 images thermiques et a obtenu une précision moyenne de 97 % [8]. Un autre modèle proposé par Berlincioni et al. était basé sur le



détecteur OpenPose suivi de l'application du modèle 3D Morphable Face pour affiner la détection du canthus intérieur. Le modèle a été entraîné sur des images visibles et a fonctionné à une vitesse de 9 FPS [9]. Enfin, la dernière tentative en date est celle de Lazri et al [10] qui ont travaillé sur la détection du canthus interne et des narines pour mesurer la température corporelle et la fréquence respiratoire. Leur méthode utilise le Single Shot Multibox Detector (SSD) pré-entraîné pour la détection des visages. Elle applique l'algorithme de Kazemi et Sullivan, disponible publiquement, pour détecter les points de repère et a donné de bons résultats sur 36 images thermiques avec une vitesse de 146 FPS. Les tentatives précédentes utilisaient un nombre limité d'images thermiques pour entraîner leurs modèles ou se basaient uniquement sur des images visibles. De plus, la détection du canthus interne était uniquement réalisée sur des images frontales.

En résumé, la détection de la fragilité physique et des signes d'infections virales chez les personnes âgées est cruciale pour intervenir rapidement et prévenir les conséquences mortelles. Des outils de dépistage et des capteurs, tels que ceux proposés par le projet ARPEGE, peuvent être utilisés pour évaluer la fragilité et la fatigue. Parallèlement, la détection du risque d'infection virale, notamment en mesurant la température corporelle à l'aide de caméras thermiques infrarouges, est essentielle. Plusieurs méthodes ont été proposées pour détecter la température du canthus interne, mais des améliorations sont encore possibles en termes de précision, de rapidité et de détection sur des images non frontales.

Dans cette étude, nous proposons une comparaison entre différents modèles de détection d'objets YOLOv5 : nano, petit, moyen et grand. Cette comparaison est présentée après avoir formé les différents modèles sur un ensemble de données de visages thermiques. La phase de préparation des données comprend la combinaison du jeu de données TFW en ligne et de nos propres images, ainsi que l'application de techniques d'augmentation des données. Nos modèles entraînés ont pour objectif de détecter les yeux, puis d'extraire la température du canthus intérieur. L'article débute par l'explication de la méthodologie, puis présente les résultats obtenus et leur discussion. Enfin, la conclusion est mise en évidence dans la dernière section.

II. MÉTHODOLOGIE

A. Ensemble de données de formation

L'ensemble de données est une combinaison d'images thermiques capturées dans notre laboratoire (104 images de 11 individus) et de l'ensemble de données en ligne du TFW (1723 images de 23 individus) [11]. Un total de 1827 images thermiques est obtenu afin d'avoir un ensemble de données plus large permettant d'obtenir un modèle plus généralisé.

La caméra thermique utilisée pour la collecte d'images thermiques dans notre laboratoire est la TROTEC IC060, avec

une résolution de 152x115. TROTEC fournit le logiciel ICReport, qui permet d'exporter les images au format JPG et les données de température correspondantes au format Excel xlsx. Cinq hommes et six femmes âgés de 22 à 30 ans se sont portés volontaires pour participer au processus de collecte des données en maintenant une distance de 1 m entre la caméra et le sujet et une température ambiante de 20°C ($\pm 1^\circ\text{C}$). L'expérience a été expliquée aux participants, qui ont accepté d'utiliser leurs images à des fins de recherche et de publication. La figure 2 ci-dessous montre le dispositif expérimental.

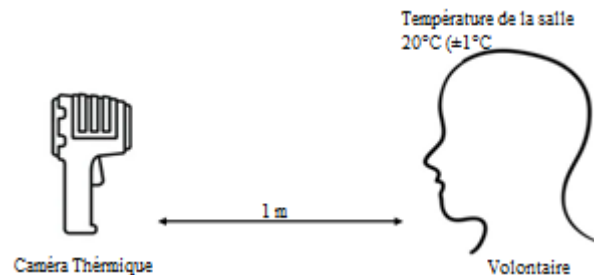


Figure 2. Configuration expérimentale de la capture d'images thermiques

Les yeux dans les images sont étiquetés à l'aide de Roboflow [12]. Deux classes sont définies : (1) "Un œil" qui inclut chaque œil dans une boîte séparée et (2) "Deux yeux" qui inclut les deux yeux dans une boîte de délimitation. La figure 3 montre un exemple d'image étiquetée avec le fichier texte d'étiquetage correspondant. Après l'étiquetage des données, le nombre d'images a été artificiellement augmenté en appliquant des techniques d'augmentation des données. Le nombre d'images est passé à 4 255 (133 %) après avoir effectué les modifications suivantes [12,13] : Retournement : horizontal ; Rotations : entre $\pm 15^\circ$; Recadrage : 0% zoom minimum à 23% zoom maximum.; Saturation : entre -25% et +25% ; Exposition : entre $\pm 5\%$; Flou.



Figure 3. Exemple d'image étiquetée et de fichier texte d'étiquettes correspondant, case rose : classe d'un œil, case jaune : classe de deux yeux ; nc = numéro de classe.

B. Modèles de formation YOLOv5

La version 5th de YOLO, YOLOv5, a été entraînée à l'aide des images thermiques préparées. YOLO est un algorithme de détection d'objets à code source ouvert basé sur le CNN. Il a été entraîné à l'origine sur l'ensemble de données COCO pour la détection de boîtes de délimitation multiples contenant des objets de différentes classes. Il calcule également le score de confiance, qui correspond à la probabilité d'avoir un objet dans une boîte englobante ou Bounding boxes en anglais, $Pr(\text{Object})$, multipliée par l'intersection ou l'union entre la boîte prédite et la boîte de vérité au sol, $(IoU_{\text{prediction}}^{\text{truth}})$, comme dans (1). Le modèle calcule également le score de confiance spécifique à la classe, qui est la probabilité que l'objet détecté appartienne à la classe i [14].

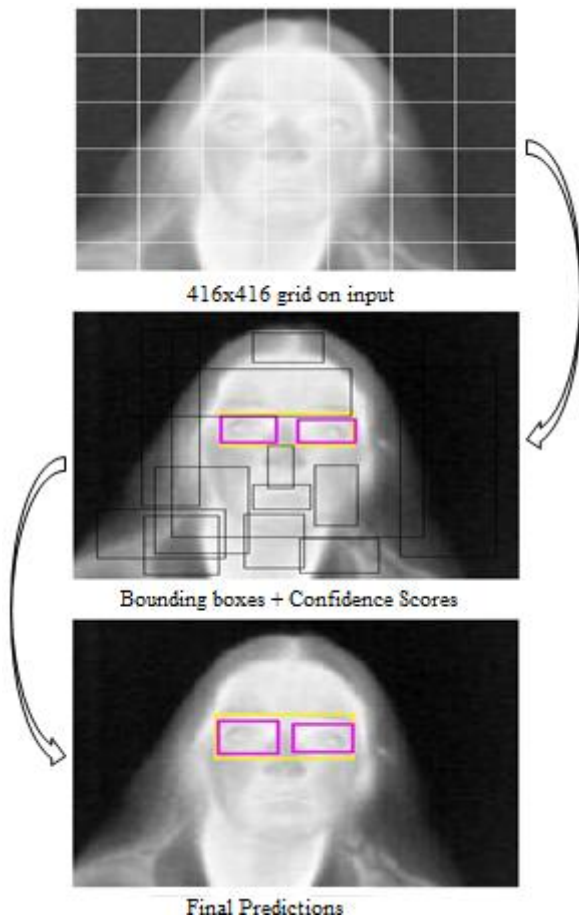


Figure 4. Les étapes de traitement du modèle entraîné sur l'image d'entrée

L'architecture du modèle YOLOv5 est composée de trois parties : l'épine dorsale (backbone), le cou (neck) et la tête (head). L'extraction des caractéristiques est effectuée par le réseau dorsal. Ensuite, la carte des caractéristiques est introduite dans le cou qui détecte les boîtes de délimitation. Enfin, la tête donne la classe, le score de confiance, l'emplacement et la taille de la ou des boîtes englobantes [15].

Dans cette étude, nous avons entraîné quatre tailles différentes de YOLOv5 : nano (n), small (s), medium (m) et large (l). Les modèles de grande taille ont une précision moyenne plus élevée mais sont plus coûteux en termes de calcul que les modèles de petite taille. Ils se distinguent les uns des autres par le nombre de couches convolutives et de sous-échantillons dans les blocs de la couche CSP de l'extracteur de caractéristiques. YOLOv5 peut améliorer les données en utilisant l'ajustement de l'espace colorimétrique, la mise à l'échelle et l'amélioration de la mosaïque [16]. Cela en fait une bonne option pour la détection des yeux dans les images thermiques. L'ensemble de données a été divisé en 80 % pour l'apprentissage et 20 % pour le test. L'apprentissage a été réalisé sur 130 époques à l'aide de la machine virtuelle Google Colaboratory, qui est un service en nuage gratuit (free cloud service) utilisé pour l'apprentissage automatique et les tâches d'analyse de données en Python [17]. Les étapes de traitement du modèle sont décrites dans la Fig 4.

C. Canthus interne Détection de la température

L'objectif d'un modèle de détection des yeux était de rechercher la température la plus élevée dans la région des yeux détectée. Opération réalisée à l'aide d'un algorithme codé sous MATLAB. Le processus se déroule comme suit :

1. Importation des coordonnées de la boîte englobante de la classe "Two Eyes
2. Importation des données excel de température qui montrent la distribution de la température dans une image
3. Début du balayage de la température la plus élevée dans les données excel, de la coordonnée du bord supérieur gauche de la boîte englobante à la coordonnée du bord inférieur droit.
4. Extraction des coordonnées de la température la plus élevée correspondant à la température du canthus interne
5. Pointer le canthus interne sur l'image thermique et imprimer sa température

III. RÉSULTATS

Les résultats de notre travail sont divisés en deux sections : (1) la performance des différents modèles formés par YOLOv5n, s, m et l, et (2) les résultats de la détection de la température du canthus intérieur.

A. Performance des modèles formés

La mesure choisie pour évaluer l'efficacité des différents modèles formés est la précision moyenne (mAP) [18]. Cependant, les mesures utilisées pour évaluer l'efficacité sont les images traitées par seconde (FPS) et la taille du modèle en mégaoctets (MB) [19]. mAP dans (2) est une mesure standard utilisée dans les applications de détection d'objets, elle est définie par l'aire sous la courbe de précision-rappel (precision-recall). Lors de l'évaluation de mAP sous la condition de $IoU \geq 0,5$, elle est donnée par $mAP@.5$.

$$mAP = \frac{\sum_{q=1}^Q AveP(q)}{Q} \quad (2)$$

où Q est le nombre de requêtes dans l'ensemble et $AveP(q)$ est la précision moyenne pour une requête donnée (q).

FPS est le nombre d'images traitées par le modèle par seconde. Des valeurs plus élevées de FPS indiquent une inférence plus rapide du modèle. Il est calculé comme indiqué dans (3), où il est la réciproque du temps d'inférence.

$$\text{FPS (frames per second)} = \frac{1}{\text{inference time(s)}} \quad (3)$$

Afin de tester l'inférence des modèles formés, une vidéo de 13 secondes a été enregistrée et utilisée. Cette vidéo montre le visage d'un volontaire regardant la caméra thermique et se tournant légèrement vers la gauche.

Les résultats des performances de YOLOv5 pour des modèles de différentes tailles sont présentés dans le tableau I.

Résultats des performances des modèles YOLOv5

Modèle	Taille du modèle (MB)	Temps d'inférence (ms)	Images traitées par seconde (FPS)	%mAP@.5
YOLOv5n	3.6	17.3	58	99.5
YOLOv5s	13.6	13.9	72	
YOLOv5m	40.1	10	100	
YOLOv5l	88.4	8.7	115 ^a	

a. FPS le plus élevé

Afin de mieux comparer les modèles, nous avons extrait la même scène de la vidéo et testé la détection des différents modèles. Par exemple, la figure 5 montre comment YOLOv5n et YOLOv5s ont détecté les yeux avec des niveaux de confiance différents. De plus, YOLOv5n a confondu à tort le chiffre "0" de la scène avec l'étiquette de la classe "OneEye".



Figure 5. La même scène vidéo avec une fausse détection de la classe "OneEye" par YOLOv5n (partie supérieure) et sans fausse détection avec YOLOv5s (partie inférieure), toutes deux avec des scores de confiance.

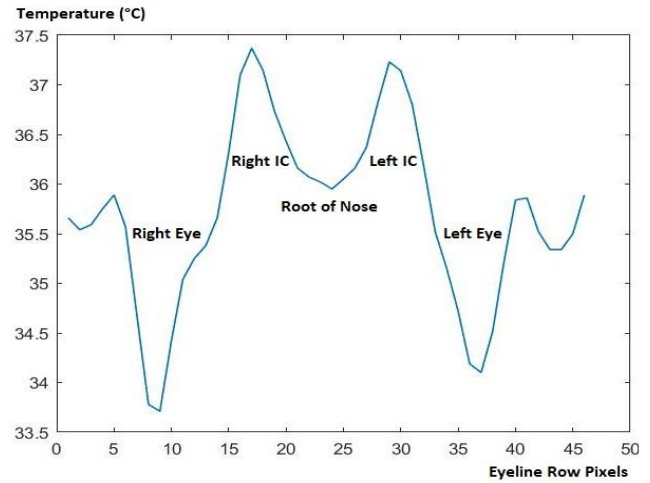


Fig6. Distribution de la température le long de la ligne des yeux, y compris : les 2 yeux, les canthus internes (CI) droit et gauche et la racine du nez.

B. Résultats de la détection de la température

La figure 6 montre une représentation graphique des températures le long de la ligne des yeux passant par les deux yeux, les deux canthus internes et la racine du nez.

La figure 7 présente un exemple de détection de la température du canthus interne dans la boîte englobante des yeux. Ce processus de détection de la température du canthus interne nécessite un temps moyen de 0,09 seconde.

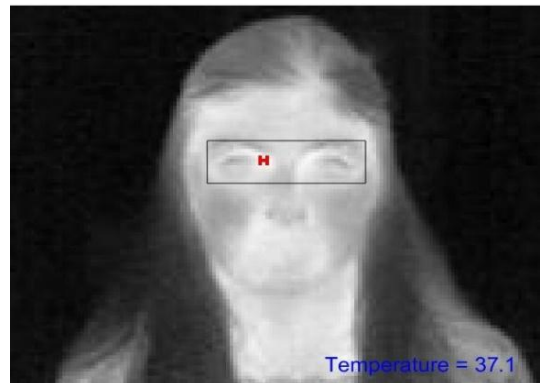


Figure 7. Exemple de détection de la température du canthus interne

IV. DISCUSSION

Les résultats du score %mAP@.5 présentés dans le tableau I étaient égaux à 99,5 % et identiques pour les différentes tailles de modèle de YOLOv5. Cependant, l'amélioration de la taille nano à la taille large a été illustrée dans le FPS qui est passé de 58 à 115. Il n'est pas surprenant que YOLOv5l ait la plus grande taille (88,4 Mo).

Les résultats de la détection dans la scène vidéo présentés dans la figure 6 prouvent que, bien que les modèles YOLOv5 aient le même %mAP@.5 global, leur efficacité et leur efficacité diffèrent lorsqu'ils traitent des cas particuliers. Dans la figure 7, les températures des deux surfaces oculaires sont

respectivement de 33,7°C et 34,1°C. Ceci est validé par l'étude de Tkáčová et al. qui indique que la température de la surface de l'œil se situe entre 34,51°C ± 0,82°C [20]. La figure 7 montre également que la température augmente de manière significative dans les régions du canthus interne droit et gauche, atteignant 37,37 °C et 37,23 °C respectivement. À la racine du nez, on observe une baisse de la température par rapport aux deux canthus internes : cela est dû à la présence de l'os nasal directement sous la peau et non des capillaires sanguins [21].

La variété de notre base de données thermiques, due à la fusion de deux ensembles de données différents et à l'application de techniques d'augmentation des données, rend notre modèle entraîné plus généralisé. Elle a également permis de créer le plus grand ensemble de données thermiques utilisé pour la détection des yeux et du canthus interne. Le deuxième plus grand ensemble de données était celui utilisé par Hussein et al. qui ont travaillé sur les caractéristiques HoG et le classificateur en cascade [22]. En outre, la présence d'images de visages inclinés dans notre ensemble de données a permis au modèle de détecter les yeux dans ces images. Cette question a été considérée comme une limitation dans d'autres études comme celle de [23].

V. CONCLUSION ET PERSPECTIVES

Le travail présenté dans cet article fournit un modèle YOLOv5 pour la détection des yeux suivi d'un algorithme pour l'extraction de la température du canthus intérieur à partir d'images thermiques. Ce système permet de comparer des modèles YOLOv5 de différentes tailles. Notre modèle pourrait être considéré comme une solution aux limitations de vitesse, de précision et de conditions d'image des méthodes de détection de la température, en particulier lors d'une épidémie.

En général, les systèmes de détection de température basés sur l'image thermique sont limités en cas de présence de lunettes en raison de leur propriété réfléchissante [9]. La solution consisterait à former un modèle qui détecte le visage et la présence de lunettes. Ensuite, le modèle recherche la température la plus élevée sur le visage plutôt que dans la région des yeux. Cela devrait permettre d'obtenir une approximation de la température du corps.

En outre, en guise de perspectives, nous visons à former et à tester YOLOv6 et YOLOv7 récemment publiés pour notre application. Cela permettra de comparer les performances des différentes versions de YOLO. En outre, notre projet étant axé sur la détection des maladies infectieuses chez les personnes âgées, nous souhaitons entraîner les modèles exclusivement sur les données thermiques des personnes âgées.

REFERENCES

- [1] Older people's health needs supported in WHO European Region, post COVID-19
- [2] Impact of aging on viral infections (nih.gov)
- [3] (2) (PDF) ARPEGE: Assessment of frailty at home (researchgate.net)
- [4] Haut Conseil de la santé publique, "Coronavirus SARS-CoV-2 : personnes à risque de formes sévères." <https://www.hcsp.fr/Explore.cgi/avisrapportsdomaine?clefr=904>
- [5] World Health Organization, "Coronavirus disease (COVID-19)," May 2021. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/question-and-answers-hub/q-a-detail/coronavirus-disease-covid-19>
- [6] R. Vardasca et al., "Bilateral assessment of body core temperature through axillar, tympanic and inner canthus thermometers in a young population," *Physiol Meas*, vol. 40, no. 9, p. 094001, Sep. 2019, doi: 10.1088/1361-6579/ab2af6.
- [7] A. Türkçetin, H. Nasibli, and M. Şahan, "Fever Detection from Human Thermal Images with Deep Learning Methods," 7TH INTERNATIONAL CONFERENCE ON ENGINEERING AND NATURAL SCIENCES (ICENS), 2021.
- [8] S. Budzan and R. Wyżgolik, "Face and eyes localization algorithm in thermal images for temperature measurement of the inner canthus of the eyes," *Infrared Phys Technol*, vol. 60, pp. 225–234, Sep. 2013, doi: 10.1016/j.infrared.2013.05.007.
- [9] C. Ferrari, L. Berlincioni, M. Bertini, and A. del Bimbo, "Inner Eye Canthus Localization for Human Body Temperature Screening," 2020 25th International Conference on Pattern Recognition (ICPR), Jan. 2021, pp. 8833–8840. doi: 10.1109/ICPR48806.2021.9412015.
- [10] Z. M. Lazri, et al, "Detecting Essential Landmarks Directly in Thermal Images for Remote Body Temperature and Respiratory Rate Measurement With a Two-Phase System," *IEEE Access*, vol. 10, pp. 39080–39094, 2022, doi: 10.1109/ACCESS.2022.3161968.
- [11] A. Kuzdeuov, D. Aubakirova, D. Koishigarina, and H. A. Varol, "TFW: Annotated Thermal Faces in the Wild Dataset," *IEEE Transactions on Information Forensics and Security*, vol. 17, pp. 2084–2094, 2022, doi: 10.1109/TIFS.2022.3177949.
- [12] "Roboflow." Available online: <https://app.roboflow.com>
- [13] C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," *J Big Data*, vol. 6, no. 1, p. 60, Dec. 2019, doi: 10.1007/s12065-019-00001-0.
- [14] D. Snegireva and A. Perkova, "Traffic Sign Recognition Application Using Yolov5 Architecture," 2021 *International Russian Automation Conference (RusAutoCon)*, 2021, pp. 1002–1007, doi: 10.1109/RusAutoCon52004.2021.9537355.
- [15] "Train Custom Data · ultralytics/yolov5 Wiki," GitHub. Available online: <https://github.com/ultralytics/yolov5/wiki/Train-Custom-Data>
- [16] Colabatory. [Online]. Available: <https://colab.research.google.com/notebooks/intro.ipynb>
- [17] M. A. Azam et al., "Deep Learning Applied to White Light and Narrow Band Imaging Videolaryngoscopy: Toward Real-Time Laryngeal Cancer Detection," *Laryngoscope*, vol. 132, no. 9, pp. 1798–1806, Sep. 2022, doi: 10.1002/lary.29960.
- [18] I. Papadeas, L. Tsochatzidis, A. Amanatiadis, and I. Pratikakis, "Real-Time Semantic Image Segmentation with Deep Learning for Autonomous Driving: A Survey," *Applied Sciences*, vol. 11, no. 19, p. 8802, Sep. 2021, doi: 10.3390/app11198802.
- [19] M. Tkáčová and P. Foffová, "A Reference for Human Eye Surface Temperature Measurements in Diagnostic Process of Ophthalmologic Diseases," *Measurement*. 2011. Available online: https://www.measurement.sk/M2011/doc/proceedings/406_Tkacova-2.pdf
- [20] "Understanding Nasal Anatomy: Inside View," *Saint Luke's Health System*. Available online: <https://www.saintlukeskc.org/health-library/understanding-nasal-anatomy-inside-view>
- [21] M. N. Hussien, M.-H. Lye, M. F. A. Fauzi, T. C. Seong, and S. Mansor, "Comparative analysis of eyes detection on face thermal images," in *2017 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*, Sep. 2017, pp. 385–389. doi: 10.1109/ICSIPA.2017.8120641.
- [22] Z. M. Lazri, et al, "Detecting Essential Landmarks Directly in Thermal Images for Remote Body Temperature and Respiratory Rate Measurement With a Two-Phase System," *IEEE Access*, vol. 10, pp. 39080–39094, 2022, doi: 10.1109/ACCESS.2022.3161968.