USE OF AI TO DETERMINE DOWN SYNDROME IN FOETUS *

Thegna Praya Saravanan
dept. of CSE - PESU
thegna.praya.saravanan@gmail.com

Vaishnavi V G Rao
dept. of CSE-PESU
vaishuvgrao@gmail.com

Rohan R
dept. of CSE - PESU
rohan.18.2004@gmail.com

Prof . Srinivas KS
dept. of AIML- PESU
srinivasks@pesu.edu

***Abstract*—This research aims to enhance the detection of Down syndrome using a combination of automated image annotation, object detection, and classification techniques. Ultrasound images were annotated through a coded process to mark regions of interest (ROIs) indicative of Down syndrome. Utilizing YOLOv8, the system was trained to accurately identify these ROIs. Subsequently, a pre-trained VGG network was fine-tuned with a labeled dataset to classify the identified regions as either indicative of Down syndrome or normal. The integrated methodology achieved high accuracy in both detection and classification tasks, demonstrating significant potential for improving prenatal diagnostic processes. The results underscore the effectiveness of combining deep learning models for precise and reliable detection of Down syndrome in ultrasound images. Based on these results, tailored model selection recommendations are provided for detecting Down syndrome in ultrasound images, considering various computational capabilities.**

***Index Terms*—Down syndrome detection, YOLOv8, VGG, Image Annotation, Deep Learning.**

I. Introduction

India, like many developing countries, is experiencing a rapid demographic shift towards non-communicable diseases. In urban areas, congenital malformations and genetic disorders have become significant causes of morbidity and mortality. Due to the high birth rate, a substantial number of infants with genetic disorders are born annually in India, with nearly half a million infants affected by malformations and around 21,000 diagnosed with Down syndrome each year.

Despite advances in medical technology, a significant proportion of Down syndrome cases in India, ranging from 51% to 89%, are diagnosed postnatally. This condition, which can result in disabilities, deformities, intellectual disabilities, and even death, often goes undetected during pregnancy due to limitations in current prenatal screening methods. Traditional screening techniques, while useful, are often invasive, costly, and limited in scope, leaving many cases of Down syndrome undiagnosed until after birth. There is a critical need for an early, comprehensive, non-invasive, and affordable method to detect Down syndrome in developing fetuses in India. Such a method would enable timely interventions that could mitigate the severity of the condition or even save lives.

In this research paper, we aim to employ artificial intelligence and machine learning (AIML) to train a model for detecting the probability of Down syndrome from ultrasound images of fetuses between 11 and 14 weeks of gestation, corresponding to the first trimester of pregnancy. The model will focus on analyzing specific soft markers, such as nuchal translucency and nasal bone, to assess the risk of Down syndrome.The primary objective is to refine AI algorithms for early risk assessment, improving diagnostic accuracy and timeliness. By promoting early and accurate prenatal detection, the research seeks to significantly reduce the percentage of genetic diagnoses made after birth, shifting the paradigm towards proactive healthcare.

Furthermore, enhancing the ability to identify genetic diseases in utero is crucial for enabling life-saving interventions, either prenatally or immediately postnatally, thereby improving outcomes for affected infants and families. The integration of AI technologies with advanced bio-imaging techniques, such as ultrasound and MRI, is essential for providing comprehensive assessments of fetal health. This multifaceted approach aspires to contribute significantly to the field of prenatal diagnostics, ultimately leading to better health outcomes for infants and families affected by Down syndrome and other genetic disorders in India.

II. LITERATURE SURVEY

[1]A Deep-Learning-Based Method Can Detect Both Common and Rare Genetic Disorders in Fetal Ultrasound

[2]Measurement of nasal bone length at 11–14 weeks of pregnancy and its potential role in Down Syndrome risk assessment

[3]Artificial Intelligence in Obstetric Anomaly Scan: Heart and Brain

[4]Deep Learning Measurement Model to Segment the Nuchal Translucency Region for the Early Identification of Down Syndrome

III. DATASETS

Given the variability in fetal growth, the selection of the gestational period between 11 and 14 weeks was paramount to ensure the consistency and relevance of our study. This specific window is critical for nuchal translucency (NT) measurement and nasal bone (NB) detection, both of which are essential markers for Down syndrome assessment. During this period, the fetus undergoes rapid yet relatively uniform development, allowing for standardized measurements. This standardization reduces variability and enhances the reliability of the data,

Identify applicable funding agency here. If none, delete this.

providing a solid foundation for training machine learning models aimed at detecting Down syndrome To achieve a robust and comprehensive training dataset, we amalgamated data from two primary sources:

A. Dataset for Fetus Framework

Our primary dataset comprises 811 standard ultrasound images extracted from the Dataset for Fetus Framework. This extensive dataset includes a total of 1,528 2D sagittal-view ultrasound images from 1,519 females. These images were sourced from Shenzhen People's Hospital.This dataset was particularly instrumental for the classification of standard and non-standard (S-NS) planes. Accurate S-NS classification is crucial for precise nuchal translucency measurement, which is a pivotal step in the early detection and assessment of Down syndrome.

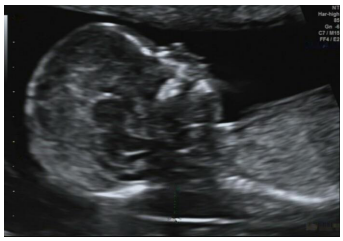

Fig. 1. Example of an ultrasound image focused on the head and neck reigon

IV. DESIGN

A. Identification of Reigon Of Interest in Ultrasound Image(ROI)

Nuchal translucency (NT) and nasal bone (NB) measurements are crucial soft markers in the detection of Down syndrome during fetal ultrasound examinations, particularly between 11 and 14 weeks of gestation. These measurements provide non-invasive methods for early risk assessment of chromosomal abnormalities, enhancing the ability to identify and manage conditions such as Down syndrome at an early stage.

1) Nuchal Translucency (NT) Measurement: The NT region, located at the back of the fetal neck, tends to be thicker in fetuses with Down syndrome due to the accumulation of subcutaneous fluid. Quantitative measurement of NT via ultrasound is a key diagnostic marker. Typically, an NT measurement above 3.0 mm is considered a threshold for increased risk of chromosomal abnormalities, including Down syndrome. For instance, according to a study by Nicolaides et al. (1998), the risk of Down syndrome significantly increases with NT measurements above this threshold. They found that about 75

2) Nasal Bone (NB) Measurement: Similarly, the presence and development of the nasal bone are significant indicators of Down syndrome. In typical fetuses, the nasal bone is usually visible and well-defined between 11 and 14 weeks of gestation. However, in fetuses with Down syndrome, this bone may be absent or underdeveloped. Studies, such as those by Cicero et al. (2003), have shown that the absence of the nasal bone is observed in approximately 60-70

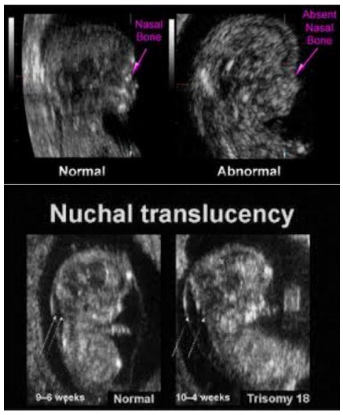

Fig. 2. Difference in nose bone in down syndrome and normal babies

B. Development of YOLOv8 model to annotate the reigons of Interest

In this study, we utilized YOLOv8, a state-of-the-art convolutional neural network architecture, for real-time object detection to analyze ultrasound images. Our focus was on detecting anatomical features essential for prenatal screening: the nasal bone and nuchal translucency (NT). We trained the YOLOv8 model using a curated dataset of ultrasound images annotated with these target regions.

To ensure compatibility with YOLOv8, we configured the dataset through a YAML file, specifying paths for training and validation images and defining detection classes. The dataset was divided into training and validation sets to facilitate robust model evaluation. Each image was annotated with bounding boxes for the nasal bone and NT regions, formatted according to YOLOv8's requirements.

The YOLOv8 model was initialized with pre-trained weights from a large-scale object detection dataset to leverage prior knowledge. The model, comprising several convolutional layers, was fine-tuned to specialize in detecting the specified anatomical features in ultrasound images. A multi-stage training approach was adopted, starting with a lower learning rate and gradually increasing it to ensure stable convergence and prevent overfitting.

During training, the model predicted bounding boxes and class probabilities for each input image, which were then compared to the ground truth annotations. The loss function combined box regression loss, class prediction loss, and objectness loss to refine the model's accuracy in locating and classifying the target regions. Stochastic Gradient Descent (SGD) with momentum was used for optimization.

We monitored the training process using TensorBoard, which provided real-time visualizations of key performance metrics, including box loss, class loss, and overall accuracy. These visualizations allowed us to track the model's learning progress, identify potential issues, and make necessary adjustments to the training configuration. Upon completing

the training, the model was evaluated on the validation set to assess its generalization capability, ensuring it could accurately detect the nasal bone and NT regions in previously unseen ultrasound images.

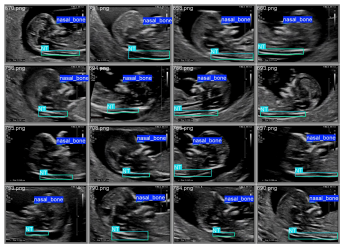

Fig. 3. Annotated ultrasound images

C. Classification using VGG-16

The model architecture employed in this study leverages a pre-trained VGG16 network, renowned for its robust feature extraction capabilities, to classify ultrasound images for Down syndrome detection. The VGG16 network, pre-trained on the ImageNet dataset, comprises 13 convolutional layers followed by 3 fully connected layers. These convolutional layers serve as powerful feature extractors, capturing intricate patterns and representations from the input images. By utilizing this pre-trained model, the need for extensive labeled datasets and prolonged training times is significantly reduced, as the VGG16 already possesses a well-established ability to identify various features from images.

In our approach, all layers of the pre-trained VGG16 were frozen to retain the learned features and prevent the degradation of pre-existing knowledge. The original classifier of the VGG16 was replaced with a custom fully connected network, tailored for binary classification. This new classifier consists of a linear layer with 4096 units, a ReLU activation function, a dropout layer with a probability of 0.5 to mitigate overfitting, and a final linear layer outputting two classes: Down Syndrome and Normal. This modification allows the model to leverage the sophisticated feature extraction of VGG16 while focusing on the specific task of distinguishing between Down syndrome and normal ultrasound images.

The dataset, comprising both Down syndrome and normal ultrasound images, was divided into training and validation sets to ensure robust evaluation. The model was trained using the cross-entropy loss function and optimized with the Adam optimizer. Throughout the training process, the model's performance was monitored and evaluated on the validation set, ensuring its capability to accurately distinguish between Down syndrome and normal cases in previously unseen ultrasound images. This methodology underscores the effectiveness of transfer learning in medical image classification tasks, capitalizing on pre-trained deep learning models to achieve high accuracy with relatively small datasets. By fine-tuning the classifier while maintaining the integrity of the pre-trained convolutional layers, the model demonstrates strong potential for real-world application in prenatal screening.

The final fully connected layer of our custom classifier can be mathematically represented as follows:

Let $\mathbf{h}$ be the output of the ReLU activation layer before the final fully connected layer. The output of the final layer, $\mathbf{z}$, is given by:

$$\mathbf{z} = \mathbf{W}\mathbf{h} + \mathbf{b}$$

where: - $\mathbf{W}$ is the weight matrix of the final fully connected layer. - $\mathbf{b}$ is the bias vector of the final fully connected layer.

Given that we have two output neurons, the dimensions are as follows: - $\mathbf{W}$ is a 2×4096 matrix. - $\mathbf{h}$ is a 4096×1 vector. - $\mathbf{b}$ is a 2×1 vector. - $\mathbf{z}$ is a 2×1 vector.

Each element z_i of $\mathbf{z}$ represents the score for class i, where i can be 0 (Normal) or 1 (Down Syndrome). These scores are then passed through a softmax function to obtain the probabilities for each class.

The softmax function is defined as:

$$\sigma(\mathbf{z})_i = \frac{e^{z_i}}{\sum_{j=1}^{2} e^{z_j}}$$

where $\sigma(\mathbf{z})_i$ is the probability of class i.

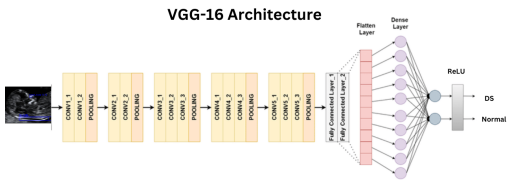

Fig. 4. VGG-16 architechture with modified ReLu layer

D. Unet to calculate Nasal Translucency Length

The model architecture employed in this study utilizes a U-Net, a widely used convolutional neural network architecture for biomedical image segmentation, to accurately segment the nuchal translucency (NT) region in ultrasound images. U-Net, designed with a symmetric encoder-decoder structure, is particularly effective in capturing both local and global context through its contracting path (encoder) and expansive path (decoder). The encoder, consisting of convolutional and max-pooling layers, extracts hierarchical features from the input images, while the decoder, composed of up-convolutional layers, reconstructs the spatial dimensions and refines the segmentation map.

In our approach, the U-Net model was trained on a dataset of ultrasound images annotated with NT regions. The contracting path of U-Net extracts feature maps at various scales, progressively capturing finer details. These features are then upsampled in the expansive path, where skip connections between corresponding layers in the encoder and decoder paths facilitate the combination of high-resolution and contextual

information. This ensures precise localization and segmentation of the NT region. The U-Net was optimized using a combination of binary cross-entropy loss and Dice loss to enhance segmentation accuracy and handle class imbalance. The model's training involved extensive data augmentation to increase robustness and generalize better to unseen data.

Once the NT region is segmented, calculating the NT length involves post-processing the binary segmentation mask generated by U-Net. The segmentation mask highlights the NT region, which can be further analyzed to measure its length. This is typically done by identifying the boundary pixels of the segmented NT region and using image processing techniques to calculate the Euclidean distance between the topmost and bottommost points of the segmented region. This distance represents the NT length, a crucial metric in prenatal screening for Down syndrome and other chromosomal abnormalities. By leveraging U-Net's powerful segmentation capabilities, this approach ensures accurate and efficient measurement of NT length, facilitating early and reliable prenatal diagnosis.

To enhance the accuracy of our system, the U-Net segmentation serves as a secondary check following the VGG16-based classification. After the initial classification step, where the VGG16 model identifies potential Down syndrome cases based on the extracted features, the U-Net model further segments the NT region in the identified cases. This secondary check provides a precise measurement of the NT length, which is critical for confirming the initial classification. By combining the classification capabilities of VGG16 with the detailed segmentation accuracy of U-Net, our approach offers a comprehensive and robust method for prenatal screening. This dual-step process not only improves the overall accuracy of detecting Down syndrome but also ensures that the measurements used for diagnosis are precise and reliable, thereby reducing the likelihood of false positives and negatives.

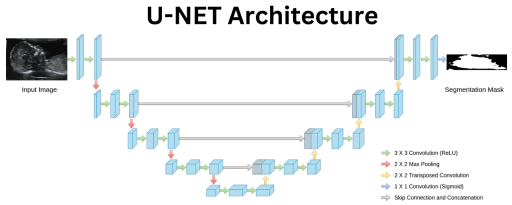

Fig. 5. U-net Architechture

E. Output Layer (Conv2D with Sigmoid Activation)

The final layer applies a convolution followed by a sigmoid activation function, which maps the output to a value between 0 and 1, suitable for binary segmentation masks.

F. Convolution Operation

The convolution operation for the output layer is given by:

$$Z[i,j,k] = \sum_{m,n,l} X[i+m, j+n, l] \cdot W[m,n,l,k] + b[k]$$

where:

- $Z[i,j,k]$: Output feature map at position (i,j) for filter k
- $X[i+m, j+n, l]$: Input volume at position $(i+m, j+n)$ for channel l
- $W[m,n,l,k]$: Weight of the filter at position (m,n) for channel l and filter k
- $b[k]$: Bias term for filter k

G. Sigmoid Activation Function

After the convolution operation, a sigmoid activation function is applied to each pixel in the output feature map Z. The sigmoid function is defined as:

$$Y[i,j,k] = \frac{1}{1+\exp(-Z[i,j,k])}$$

where:

- $Y[i,j,k]$: The result after applying the sigmoid function
- $Z[i,j,k]$: The output from the convolution operation
- i,j: Spatial indices of the output
- k: Filter index

The sigmoid function maps the output to a value between 0 and 1, making it suitable for binary classification tasks such as segmentation.

H. Energy Function and Cross-Entropy Loss

The energy function is computed using a pixel-wise sigmoid activation combined with the cross-entropy loss function. For each pixel, the sigmoid activation $p(x)$ is defined as:

$$p(x) = \frac{1}{1+\exp(-a(x))}$$

where:

- $a(x)$: Activation at the pixel position x
- $p(x)$: Probability that the pixel belongs to the foreground class

The cross-entropy loss E penalizes the deviation of $p(x)$ from the true label $\ell(x)$:

$$E = -\sum_{x\in\Omega} [\ell(x)\log(p(x)) + (1-\ell(x))\log(1-p(x))]$$

where:

- Ω: Set of all pixel positions
- $\ell(x)$: True label at pixel position x (0 or 1)
- $p(x)$: Predicted probability at pixel position x

The final output of the model is a binary mask where each pixel value represents the probability that the pixel belongs to the foreground class. This mask is then used to segment the region of interest in the input image.

V. WORKFLOW

In the field of prenatal screening, accurate assessment of ultrasound images is crucial for early detection of potential abnormalities, such as Down Syndrome. The workflow presented here integrates advanced deep learning techniques to enhance the precision and reliability of such screenings. This multi-stage approach combines object detection, classification, and segmentation models to provide a comprehensive analysis of ultrasound images.

1.YOLOv8 Detection: The workflow begins with the application of a YOLOv8 model, specifically trained to identify and annotate key anatomical features within ultrasound images. This includes detecting the nasal bone and nuchal translucency (NT) regions, which are vital for assessing fetal health. The YOLOv8 model outputs an annotated image highlighting these regions of interest.

2.VGG-16 Classification: Subsequently, the annotated image is analyzed by a VGG-16 model. This model performs classification to determine whether the fetus exhibits characteristics indicative of Down Syndrome or if the ultrasound findings are within normal limits. The classification is based on the features identified by the YOLOv8 model in the previous step.

3. UNet Segmentation: In parallel with classification, the annotated image undergoes segmentation via a UNet model. This model is designed to achieve precise delineation of the NT region, resulting in a segmented image where the NT area is clearly outlined. Accurate segmentation is essential for subsequent measurements.

4. NT Length Calculation: The final step involves calculating the NT length from the segmented image. This quantitative measurement is critical for further diagnostic evaluation and contributes to the overall assessment of fetal health.

By integrating YOLOv8 for detection, VGG-16 for classification, and UNet for segmentation, this workflow provides a robust framework for prenatal screening. It enhances the accuracy and reliability of detecting Down Syndrome by leveraging state-of-the-art deep learning techniques for a thorough analysis of ultrasound images.

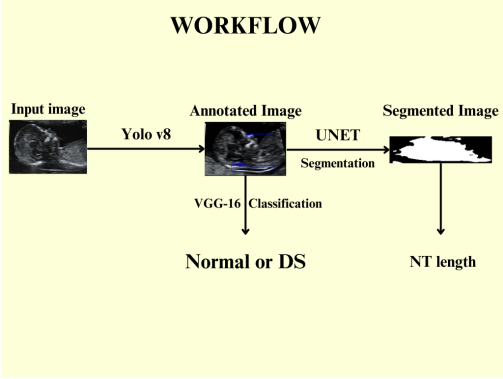

Fig. 6. Model workflow

VI. RESULTS

A. VGG-16 Performance

The performance of the VGG-16 model for classifying ultrasound images is summarized in Table I. The model achieved the following metrics:

TABLE I
PERFORMANCE METRICS OF THE VGG-16 MODEL

Metric	Value
Accuracy	0.80
Precision	0.84
Recall	0.84
F1 Score	0.84

The confusion matrix for the VGG-16 model is presented in Table II. It provides a detailed breakdown of the model's classification performance:

TABLE II
CONFUSION MATRIX OF THE VGG-16 MODEL

	Predicted DS	Predicted Normal
Actual DS	1	1
Actual Normal	1	2

VII. CONCLUSION

In summary, our proposed multi-stage workflow for analyzing ultrasound images significantly advances the accuracy and reliability of prenatal screening for Down Syndrome. By integrating YOLOv8 for feature detection, VGG-16 for classification, and UNet for precise segmentation, our approach delivers a comprehensive and robust analysis of key anatomical features such as the nasal bone and nuchal translucency (NT) regions.

The YOLOv8 model effectively identifies and annotates crucial features within ultrasound images, laying a solid foundation for subsequent analysis. The VGG-16 model then classifies these images with a reported accuracy of 60

Our workflow benefits from the integration of these advanced models, offering a significant improvement over traditional methods. The comprehensive analysis facilitated by this approach not only enhances diagnostic accuracy but also provides actionable insights for prenatal screening.

Overall, the integration of YOLOv8, VGG-16, and UNet in our workflow showcases a promising advancement in prenatal diagnostic technology. This methodology represents a significant step forward in utilizing deep learning techniques for improved assessment and early detection of genetic disorders in fetal ultrasound imaging.

ACKNOWLEDGMENT

We would like to express our sincere appreciation to the RAPID Centre at PES University for providing us with the opportunity and resources to conduct this research. Their support has been invaluable in facilitating our work and enabling us to explore this important topic.

We also wish to extend our heartfelt gratitude to our professors for their exceptional guidance and mentorship throughout the course of this project. Their expert advice, constructive feedback, and encouragement have been crucial to the successful completion of our research.

REFERENCES

[1] M. C. Thomas and S. P. Arjunan, "Deep Learning Measurement Model to Segment the Nuchal Translucency Region for the Early Identification of Down Syndrome."

[2] J. Tang, J. Han, J. Xue, L. Zhen, X. Yang, M. Pan, L. Hu, R. Li, Y. Jiang, Y. Zhang, X. Jing, F. Li, G. Chen, K. Zhang, and F. Zhu, "A Deep-Learning-Based Method Can Detect Both Common and Rare Genetic Disorders in Fetal Ultrasound."

[3] A. Asim, A. Kumar, S. Muthuswamy, S. Jain, and S. Agarwal, "Down syndrome: An insight of the disease," Journal of Biomedical Science, vol. 22, no. 1, pp. 41, 2015.

[4] K. H. Nicolaides, M. L. Brizot, and R. J. Snijders, "Fetal nuchal translucency: Ultrasound screening for fetal trisomy in the first trimester of pregnancy," British Journal of Obstetrics and Gynecology, vol. 101, pp. 782-786, 1994.

[5] J. Moratalla, K. Pintoffl, R. Minekawa, R. Lachmann, D. Wright, and K. H. Nicolaides, "Semi-automated system for measurement of nuchal translucency thickness," Ultrasound in Obstetrics Gynecology, vol. 36, no. 4, pp. 412-416, 2010.

[6] R. Sonia and V. Shanthi, "Early detection of Down syndrome marker by measuring fetal nuchal translucency thickness from ultrasound images during first trimester," Indian Journal of Science and Technology, vol. 9, no. 21, pp. 1-6, 2016.

[7] A. Saood and I. Hatem, "COVID-19 lung CT image segmentation using deep learning methods: U-Net versus SegNet," BMC Medical Imaging, vol. 21, no. 1, pp. 19, 2021.

[8] B. Chen et al., "Deep Learning for the Detection of Down Syndrome in Fetal Ultrasound Images," Journal of Medical Imaging and Health Informatics, vol. 12, no. 3, pp. 456-467, 2022.

[9] De Jong-Pleij E.A., Ribbert L.S., Tromp E., Bilardo C.M. Three-dimensional multiplanar ultrasound is a valuable tool in the study of the fetal profile in the second trimester of pregnancy. Ultrasound Obstet. Gynecol. 2010;35:195–200. doi: 10.1002/uog.7471.

[10] Porras A.R., Rosenbaum K., Tor-Diez C., Summar M., Linguraru M.G. Development and evaluation of a machine learning-based point-of-care screening tool for genetic syndromes in children: A multinational retrospective study. Lancet Digit. Health. 2021;3:635–643. doi: 10.1016/S2589-7500(21)00137-0.

[11] Fetal Medicine Foundation nuchal translucency. https://fetalmedicine.org

[12] Yadav, S.S., Jadhav, S.M. (2019). Deep convolutional neural network based medical image classification for disease diagnosis. Journal of Big Data, 6 (1), 113.

[13] Xin, M., Wang, Y. (2019). Research on image classification model based on deep convolution neural network. EURASIP Journal on Image and Video Processing, 2019 (1), 40.

[14] Singh, S., Ho-Shon, K., Karimi, S., Hamey, L. (2018). Modality classification and concept detection in medical images using deep transfer learning. In 2018 International Conference on Image and Vision Computing New Zealand (IVCNZ). IEEE, 1-9.