Wavelets Final Project Report
EE678

Mohit
20D070052

November 2023

1 Student Details

Name:	Mohit
Roll No:	20D070052
Group No:	7
Group Member:	Vinayak Goyal (20D070088)

2 Problem Statement

It is well recognised that wavelets and multiresolution techniques provide appealing and efficient solutions to address the issue of biometric information verification, identification, and authentication.

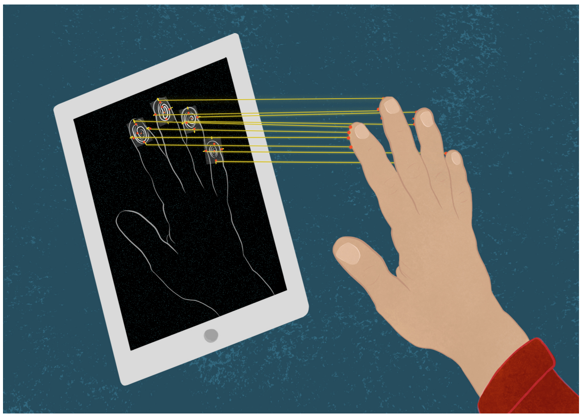
For this course project, we will only be working on physiological biometrics . These consist of fingerprints and iris scans in the beginning, followed by other biometrics like the ear, face, knuckles, and so on, which could be used in future.

Define and clarify what is meant by identification, authentication, and verification in relation to different physiological biometrics that define ”figures of merit” for assessment.

Regarding fingerprints: Determine the most efficient way to depict fingerprints using wavelets. As Dr. Azhar explained, start with separable techniques, which are basically ‘tensor product’ approaches with one-dimensional wavelets. To complete the duties of identification, verification, and authentication, use these representations. Do this while introducing and studying as many potential distortions as you can. Also machine learning methodologies, might be utilised conjunction with the “scattering wavelet network” methodology developed by Stephane Mallat. Introduce strategies to enhance your performance, such as potentially using multiple fingerprints from distinct fingers for the same person or people, rather than just one.

In fact, multiresolution and non-separable wavelet techniques offer appealing biometrics possibilities. Shearlet and curvelet techniques are two examples. Nonetheless, a lot of researchers have a tendency to employ non-separable methodologies, either without fully utilising their benefits or without considering the possibility of combining the finest features of both separable and non-separable approaches. Clearly, combining the two of these ideas will only improve things. Show how to accomplish this with both iris and fingerprints. It will be necessary to do this by deftly using non-separable techniques, maybe in addition to separable ones. Without such expertise and understanding, merely using non-separable procedures is unlikely to provide positive outcomes.

Provide conceptual, analytical, or theoretical ideas that have surfaced during the course of your research and implementations to bolster your findings.



3 Motivation

This project will introduce particular field of signal processing research and development (R&D) and provide the knowledge and skills necessary to thoroughly examine a topic of contemporary significance from all perspectives: conceptual (theoretical, analytical), implementation (realisation, programming, benchmarking), and exposure.

It takes on extra significance because IIT Bombay is being considered by the Unique Identification Authority of India (UIDAI) to lead the country's development of cutting edge contactless biometric technology, which will undoubtedly benefit from strong performance in this course project.

4 Datasets Used

4.1 Dataset 1 - Fingerprint Dataset

Initially we plan to use the fingerprint dataset : PolyU Contactless Fingerprint to Contact-based Fingerprint Database. This database contains 2976 contatcless 2D fingeprint images and conresponding 2976 contact-based fingeprint from 336 clients. Six contactless and contact-based fingerprint images (impressions) were acquired from each finger.

The size of each contatcless 2D fingerprint image is around 3.60 MB. The size of each contact-based fingerprint is around 64.00 KB. The size of each downsampled contatcless 2D fingerprint image is around 79.00 KB. The database contains contatcless 2D fingerprint images which are stored in bitmap and contact-based fingerprints which are stored in JEPG format.

The first session part of database was acquired from 336 different clients/fingers. Each of the client provided 6 different fingerprint samples (6 images). The second session part of the database contains images from corresponding 160 clients, and each of these second-session clients provided 6 fingerprint samples (6 images) after an interval of 2-24 months. Therefore, there a total of 5952 images were acquired for this database. Given are some of the sample images from the dataset.

Figure 1: Contact sensor image

Figure 2: Contactless sensor image(camera)

Figure 3: Processed Contactless sensor image (grayscaled)

4.2 Dataset 2 - Iris Dataset

This iris image database mainly consists of the iris images collected from the students and staff at IIT Delhi, India. The acquired images are saved in bitmap format. The database of 2240 images is acquired from 224 different users. All the subjects in the database are in the age group 14-55 years comprising of 176 males and 48 females. The resolution of these images is 320 x 240 pixels and all these images are acquired in the indoor environment. The images were acquired using an automated program that requires users to present their eyes in a sequence until ten images are registered.

The acquired database is saved in 224 folders, each corresponding to 224 subjects. Majority of images were acquired from the left eyes while the rest images were acquired from right eye. Now the database has a label 'L' or 'R' which designates left or right eye. There are 1288 images from 224 subject that are from left eyes while the rest images from 211 subjects are from right eyes.Except folders 1-13, 27, 55 and 65 all other folders have five left and 5 right eye images.

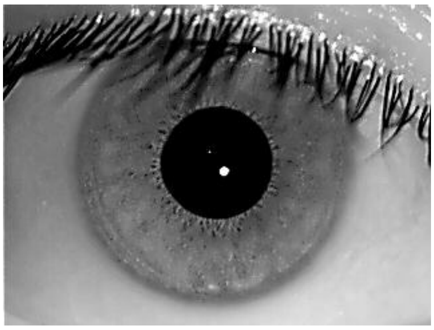


Figure 4: Example Left Eye Image from Dataset

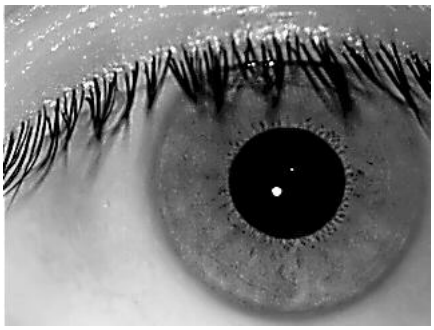


Figure 5: Example Right Eye Image from Dataset

5 Biometrics Image Pre-Processing

Preprocessing steps for fingerprint images are essential to enhance the quality of the images and prepare them for further analysis or recognition tasks. There are various preprocessing algorithms that can be used which each has there own advantages. Given below is a list of pre processing that can be used for biometrics fingerprints however we will need to try these methods before applying directly :

- **Histogram Equalization** : Histogram equalization is beneficial for fingerprint datasets because it can improve the visibility of features in the image, such as ridge patterns and minutiae points. This can make it easier for fingerprint recognition algorithms to extract and match features, leading to better recognition performance. Additionally, histogram equalization can help to normalize the brightness of fingerprint images, which can make them more robust to variations in lighting conditions.

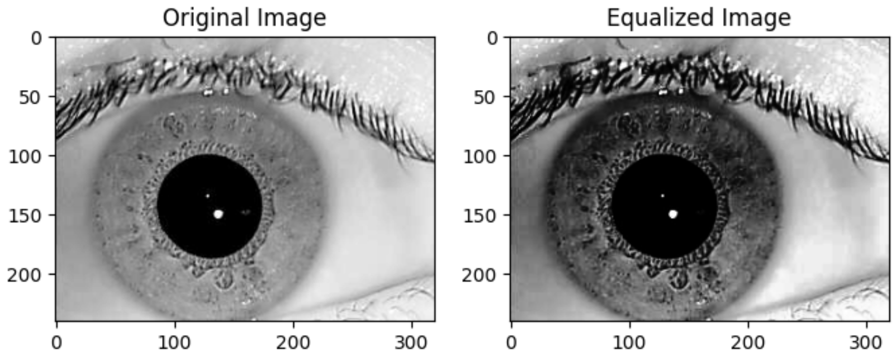


Figure 6: Visualizing the image of eye after applying the histogram equalization

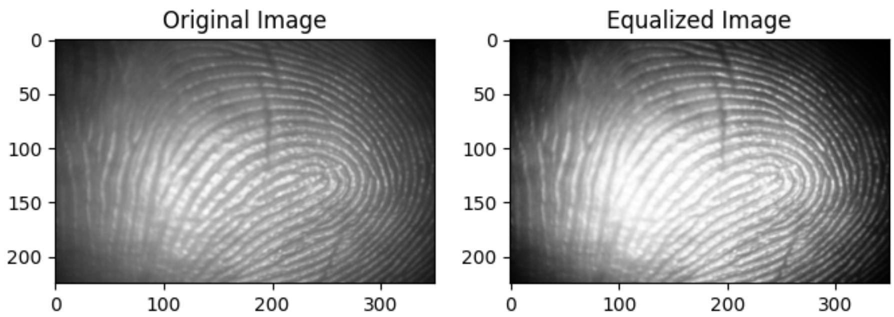


Figure 7: Visualizing the image of finger after applying the histogram equalization

- **Binarization/ Adaptive Thresholding** : Adaptive thresholding is a type of binarization that uses a different threshold value for each pixel in the image. This can help to preserve the fine details of the fingerprint, while still removing noise and artifacts.

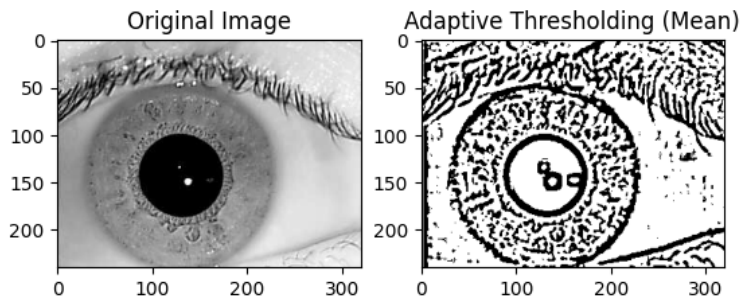


Figure 8: Visualizing the image of eye after applying the adaptive thresholding

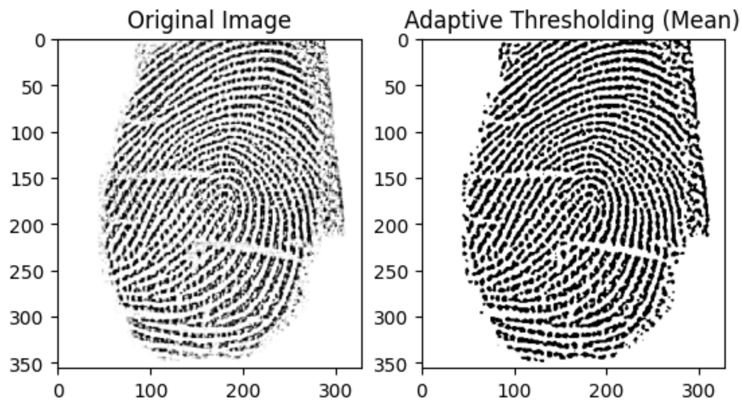


Figure 9: Visualizing the image of contact finger after applying the adaptive thresholding

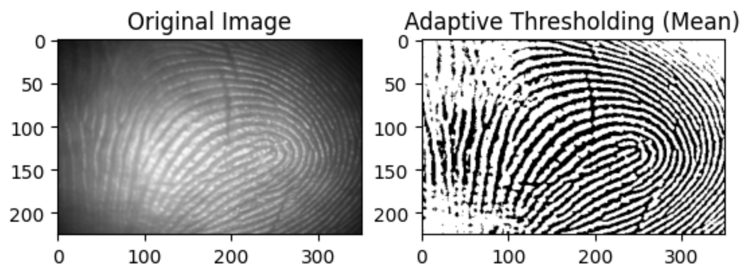


Figure 10: Visualizing the image of contactless finger after applying the adaptive thresholding

- **Minutiae Extraction** :Minutiae points are unique and stable features of fingerprints which reduces the complex fingerprint image into a set of coordinates and orientations, significantly reducing storage requirements and computational complexity. This makes it easier to compare and match fingerprints efficiently.

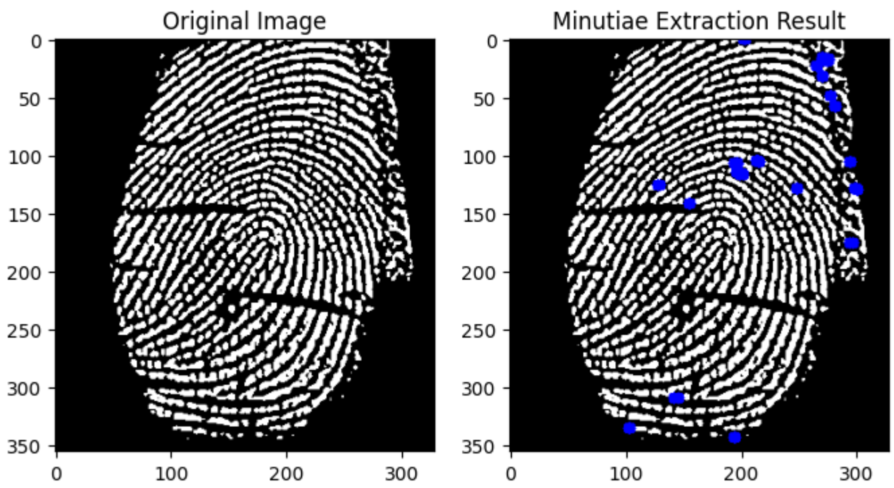


Figure 11: Visualizing the image of contactbased finger after applying the minutiae extraction

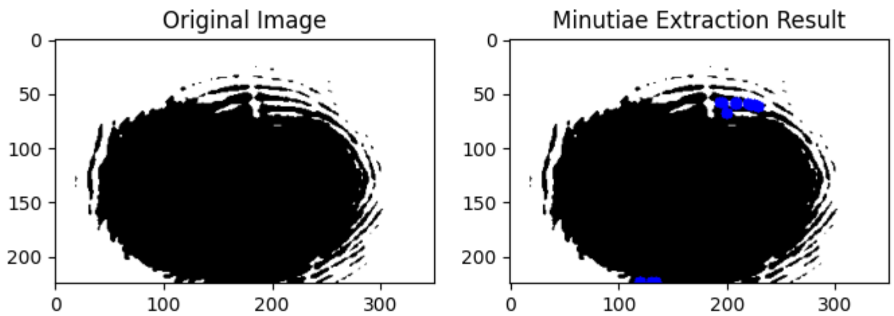


Figure 12: Visualizing the image of contact less based finger after applying the minutiae extraction

- **Sobel Operator** : It is a simple and efficient edge detector that can be implemented using a convolution kernel. The Sobel operator calculates the gradient magnitude and orientation at each pixel in the image. The gradient magnitude is a measure of how strong the edge is at that pixel, and the gradient orientation is a measure of the direction of the edge. This operator is sensitive to noise but we will already use the noise cancellation to remove the noise.

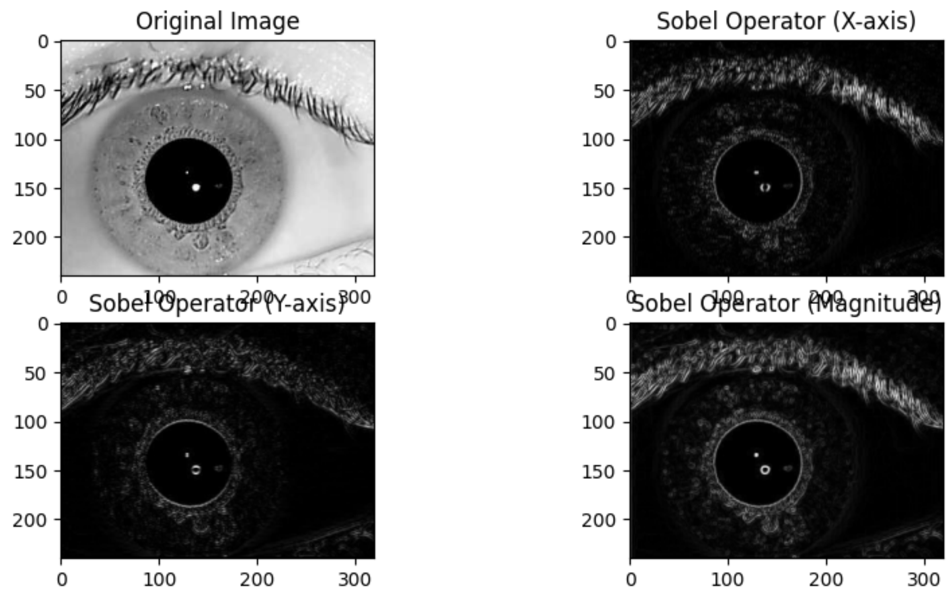


Figure 13: Visualizing the image of eye after applying the sobel operator

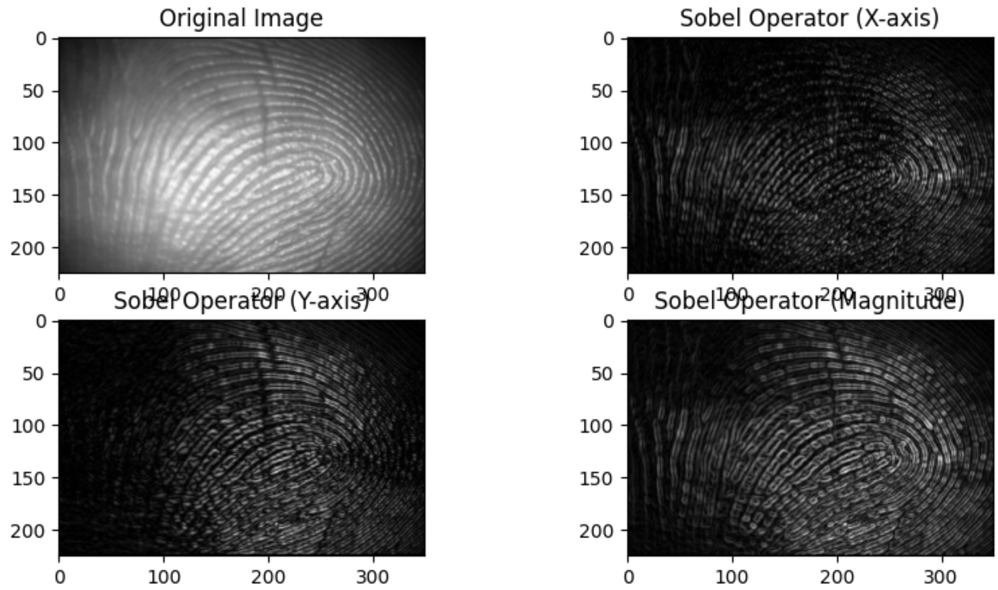


Figure 14: Visualizing the image of finger after applying the sobel operator

- **Gabor Operator** : Gabor filters are useful in preprocessing fingerprint images due to their ability to enhance ridge patterns, provide frequency and orientation selectivity, and reduce noise. They are applied to fingerprint images to highlight ridge features, making it easier to extract relevant information for fingerprint recognition. The filters are particularly effective in capturing variations in intensity along ridges and can be tuned to specific frequencies and orientations. Gabor-filtered images serve as a basis for feature extraction, including ridge count, width, and minutiae points, essential for fingerprint recognition systems.

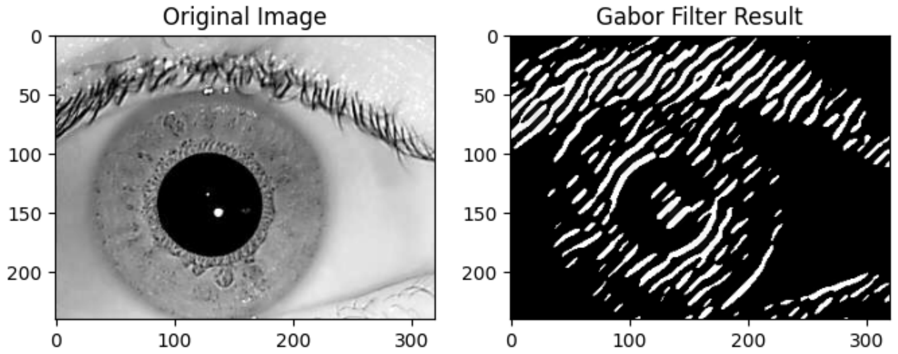


Figure 15: Visualizing the image of eye after applying the gabor filter

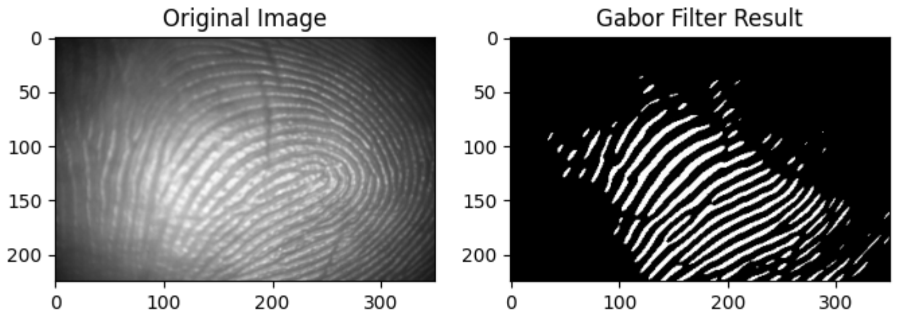


Figure 16: Visualizing the image of finger after applying the gabor filter

6 Scattering Wavelet Transform

The Scattering Wavelet Network (SWN) is valuable in biometrics verification by offering translation invariance, robustness to deformations, and multiresolution analysis. Its hierarchical structure facilitates discriminative feature extraction, enhancing accuracy. SWN's reduced dimensionality aids efficiency, and its compatibility with machine learning enables integration into advanced verification systems. Overall, SWN's ability to capture unique biometric traits, despite variations, makes it a powerful tool for reliable and robust biometric verification.

To apply the scattering wavelet transform we have used the kymatio library in python which gives the scattering coefficients upto level 2 using Morlet wavelet.

- Depth of Scattering Network (J) - In wavelet analysis, increasing the depth corresponds to analyzing the signal at coarser resolutions or lower frequencies, indicating how many times the signal has been transformed. Increasing J can lead to a large number of parameters therefore in most cases we limit ourselves to J = 2.
- Angle Index (L) - This is used as an index for spatial or angle information. For example, in the case of 2D scattering, this corresponds to different orientations in the image. In other words, it represents the spatial or angular position of the wavelet at a particular scale. This is relevant when dealing with oriented wavelets or filters, such as in the Morlet wavelet family.

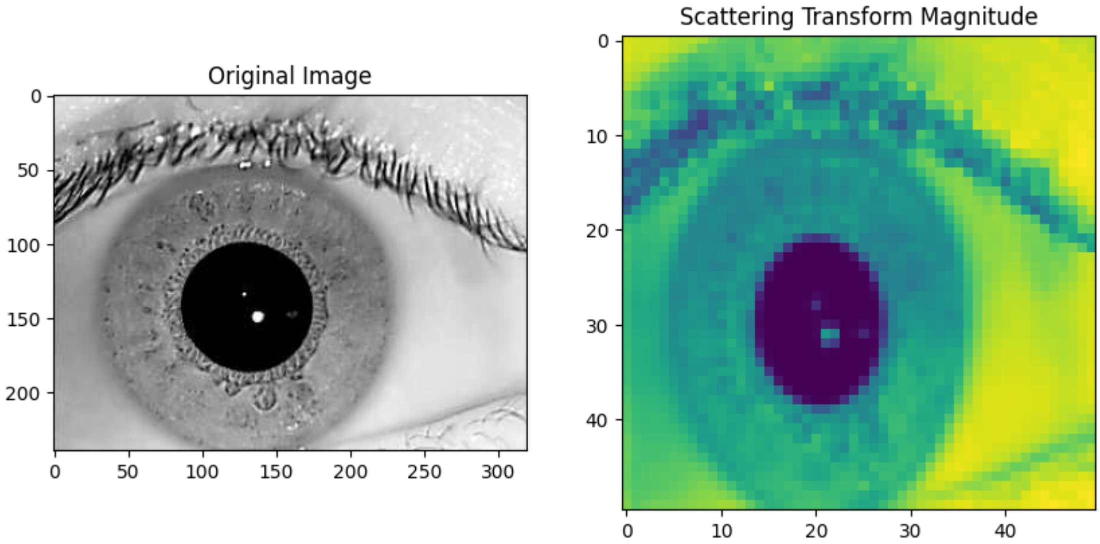


Figure 17: Visualizing the image of eye after applying the scattering wavelet transform

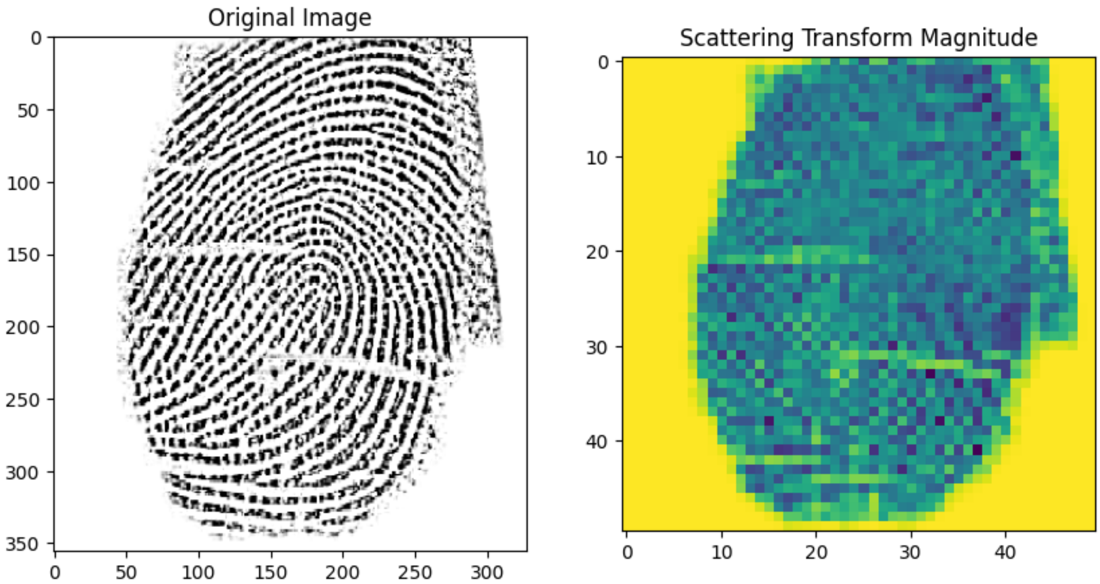


Figure 18: Visualizing the image of contact based fingerprint after applying the scattering wavelet transform

7 Machine Learning Model

In general a Machine Learning model for multi label classification is comprised of a feature extractor like resnet, efficientNet and other advanced networks with a basic classifier built upon it. However these feature extractors have a very large number of trainable parameters due to which it takes a huge time to train the model, also these are not that efficient for fingerprint images.

In our model we have used scattering wavelet transform as the feature extractor instead of other feature extractors like resnet etc. due to its translation invariance, stability to deformations, hierarchical representation capturing multiscale information, discriminative power, and compatibility with deep learning. These properties make it suitable for robustly capturing relevant fingerprint patterns, including ridges and minutiae, across various conditions and scales.

On this feature extractor we have built a custom classifier which is comprised of flattening, then dense layers and some regularizations like L2 regularization and dropout. Along with this we have used some convolutional blocks and batch normalization blocks because it has some advantages like they are used in classifiers due to their ability to perform localized feature extraction, share parameters for efficiency, provide translation invariance, enable hierarchical feature learning, reduce sensitivity to local deformations, and effectively parameterize spatial patterns. However we can't use more and more convolutional layers as it will drastically reduce the number of training parameters which leads to overfitting.

After this we have used the optimizer as ADAM as it is one of the best known optimizer and to calculate the loss we have used cross entropy which will be backpropagated to train the weights, according to these matrices we will we will calculate the accuracy of our model by varying these parameters.

The results obtained from the model is given below for various parameters :

Depth (J)	Angle Index (L)	Convolutional Blocks	Trainable Parameters	Test Accuracy
1	6	1	3,607,971	85.71
2	6	1	16,308,771	84.08
1	8	1	5,019,171	86.90
2	8	1	27,367,971	80.65
1	6	2	2,162,883	84.23
2	6	2	14,978,883	74.70
1	8	2	3,545,283	81.85
2	8	2	25,577,283	79.76

Figure 19: Results on Applying the ML model for different parameters of scattering wavelet transform and number of convolutional blocks

8 Shearlet Application

We have tried to incorporate the non-separable approach for the image feature extraction. The non-separable methods include shearlets and curvelets as taught by Dr. Azhar. However, these are by default not available in python as library therefore we needed to code the shearlet from scratch. We have successfully been able to code and get transforms of the shearlets and these are the results till now. Since our feature extraction part is completed, we will now work on identify and using these features efficiently and accurately.

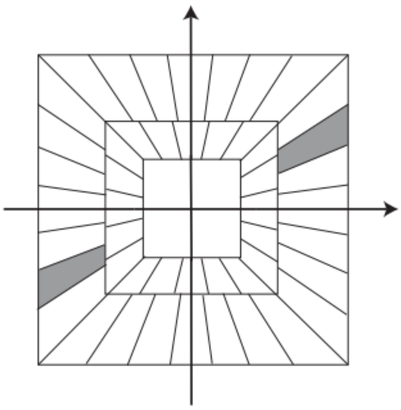


Figure 20: Tiling of Fourier domain induced by the cone-adapted discrete shearlet system associated with classical shearlets

Work done during endsem:

Further, we tried the extension of the Shearlet transform upto level 2. We hypothesised that due to implementation in python and the domain actually being discrete(image), two adjacent tiles of the cone-adapted shearlet system will have a small intersection. This small intersection can be used to get very accurate edges perpendicular to the original image.

Also, after discussion on 18th November, we decided to reduce the support of the Meyer wavelet which forms the Shearlet system. This is because of the following reasons:

1. We will be able to cover smaller frequency regions.

2. There will be more tiles covering the higher frequency region and they will be thin and contain more specific information.

3. There are total of 44 masks in level 1 shearlet transform of the new meyer wavelet system. So it seems that taking one mask and applying another mask on it will give a total of 44*44= 1936 masks, but most of the masks do not have any region in common and thus their combination accounts to zero. Thus we calculated the **energy** in a mask saved only those with non-zero energy. This lead to a total of 208 masks only in level 2 shearlet transform.

Our new Meyer wavelet has support of $[-\frac{1}{16}, -\frac{1}{64}]U[\frac{1}{64}, \frac{1}{16}]$

We calculated the cosine similarity using different ways. We took the shearlet transform and calculated the Euclidean distance for each of the outputs. Then we took a mean to find the average Eucledian distance. We wanted to see if the average Euclidean distance for direct Shearlet transform is less for same fingerprint or not. However, the distance was similar for all people. We also used different fingers for a single person and came to a similar conclusion that the Eucledian distance without processing is not a good metric.

Figure 21: Eucledian distance calculated between Shearlet transform of 1_1 (person1_fingerprint1) with all six fingerprints of 20 people; first row person one; second row person two and so on

Cosine distance of 1_1.jpg from all 6 images of person 1 to 20

0.0	917.9490833257136	1017.8322466696192	990.7118783956026	1059.156973734268	991.2830361138098
980.8942874495514	1005.8267178923917	995.4654142592216	985.7321546931363	986.8573233230301	987.5362373816856
1062.93397431014	1071.879530003698	1040.2714237843966	990.0898202916349	964.2963958690555	1053.693099874498
1000.5637377448598	1019.6887158775978	990.6283571302712	1037.2276521540646	1013.0028455170042	1016.8922767930723
1003.8688962636575	1012.3536266359133	1021.2874020899711	1000.8817563273839	1027.9482408374502	1012.8896917633622
1021.5065826903599	998.959194499863	996.73040503535	1017.3717475075478	1021.2944422511621	1026.8078502624937
1037.2076918971788	1009.842076973707	1028.5350555587595	1036.2997070519461	1027.6081970239413	1039.8132090360402
997.3450982084282	1005.3709686599097	1007.2731392989889	1001.0990351502268	1012.3189988859209	1007.069745685565
979.9058601306486	1011.3546894801476	970.996988842557	995.3561518518538	961.8820769420024	981.5206715216991
1018.6626800219871	992.2501544759925	1001.8062676328325	997.662806007772	1006.0192241109758	991.297452068801
984.3202198404572	978.3503976943372	1017.514534027947	985.9121685635887	978.6111992438025	963.7554361577834
931.2222973540312	956.3337880213239	932.0486174798309	912.9933465039499	961.994116598197	957.5862537026347
981.9953075411481	961.8493681548784	959.0559868868371	994.4664833636234	995.4450725693976	956.607446193627
995.9384484680759	1029.2864037061752	1031.0553049062935	1058.919706682046	1031.7949084510053	1052.3174471230514
1029.620979948619	1038.8032551704719	1004.2527417596613	1063.9493989121356	1055.3657697137069	987.6938388201062
957.1463117215536	947.3430412239388	987.4908828902636	998.6180018385971	1014.7615372264564	978.6557765223806
1036.9101798776524	1070.8724229700579	1036.018814008234	1062.454250834044	1017.1371470715702	1007.7051088231727
1024.1401504304624	1038.3617642667368	998.2260522364381	1017.691473318385	1019.5584941759348	981.900258332499
1047.602307682059	1045.5041079341295	1017.0141875416393	1012.1643717761071	1059.0765118359739	1061.3236491642256
1001.2600333500918	989.1846382626019	984.8601594028968	1003.0249028157546	1005.3981288851005	982.4184694975374

We also tried using these 208 features in a classification model. We will also try training a Siamese network for this but right now we do not have results in CNN or Siamese network as model had many parameters and took a lot of time to train. We will try showing these results.

I have shown the masks of the level 1 and level 2 shearlet below:

Figure 22: Filter mask and filtered fingerprint for values of j and s for both ψ and $\tilde{\psi}$ of a scanned fingerprint

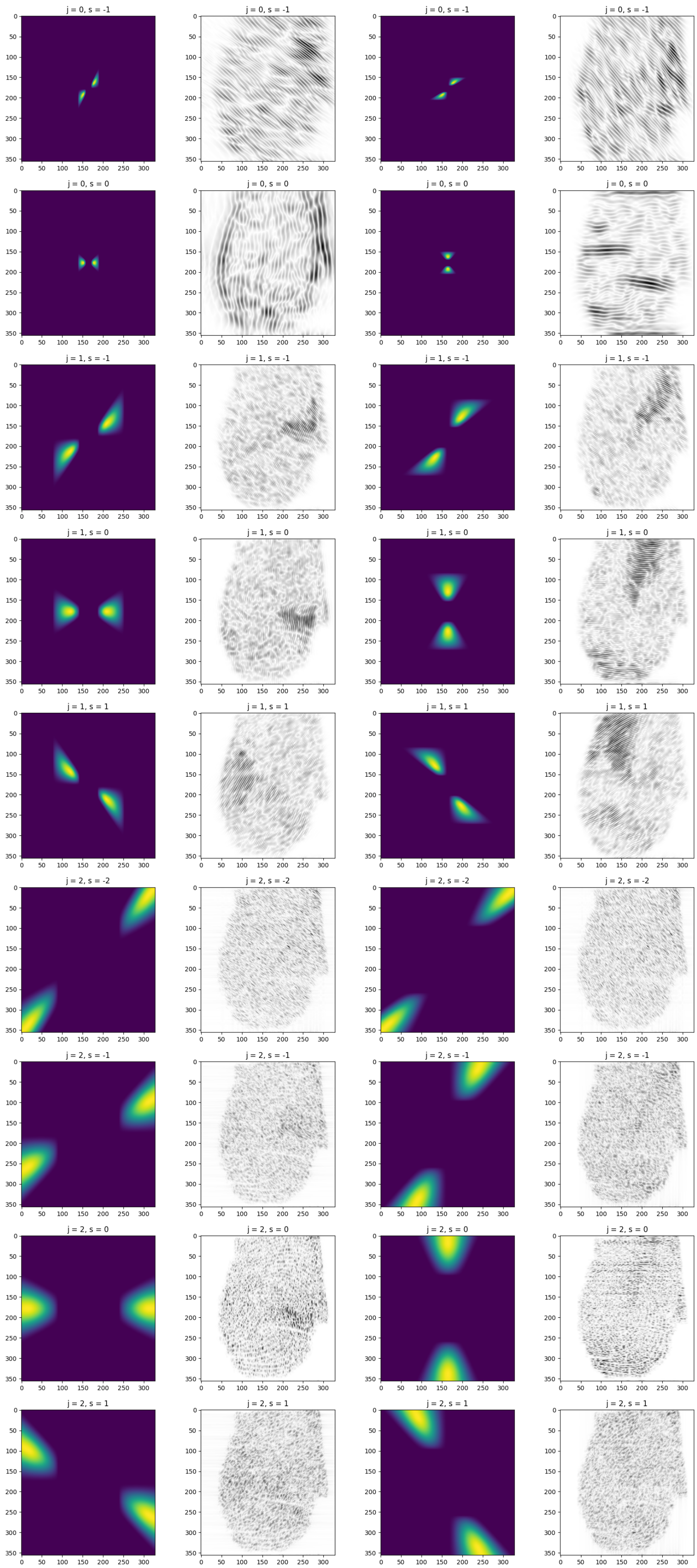


Figure 23: Level 1 Masks of new Meyer wavelet

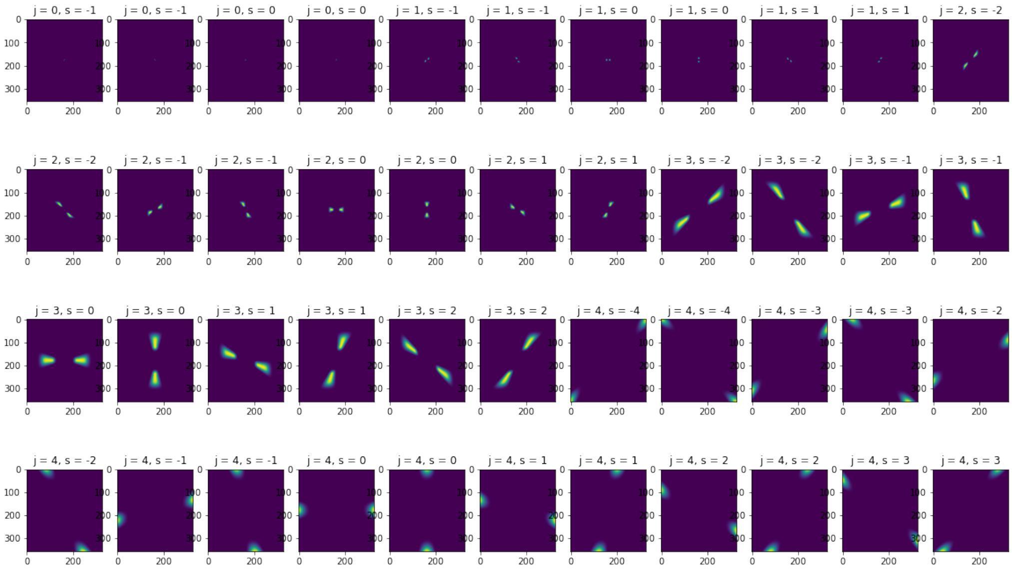


Figure 24: Output of new level 1 masks on a fingerprint

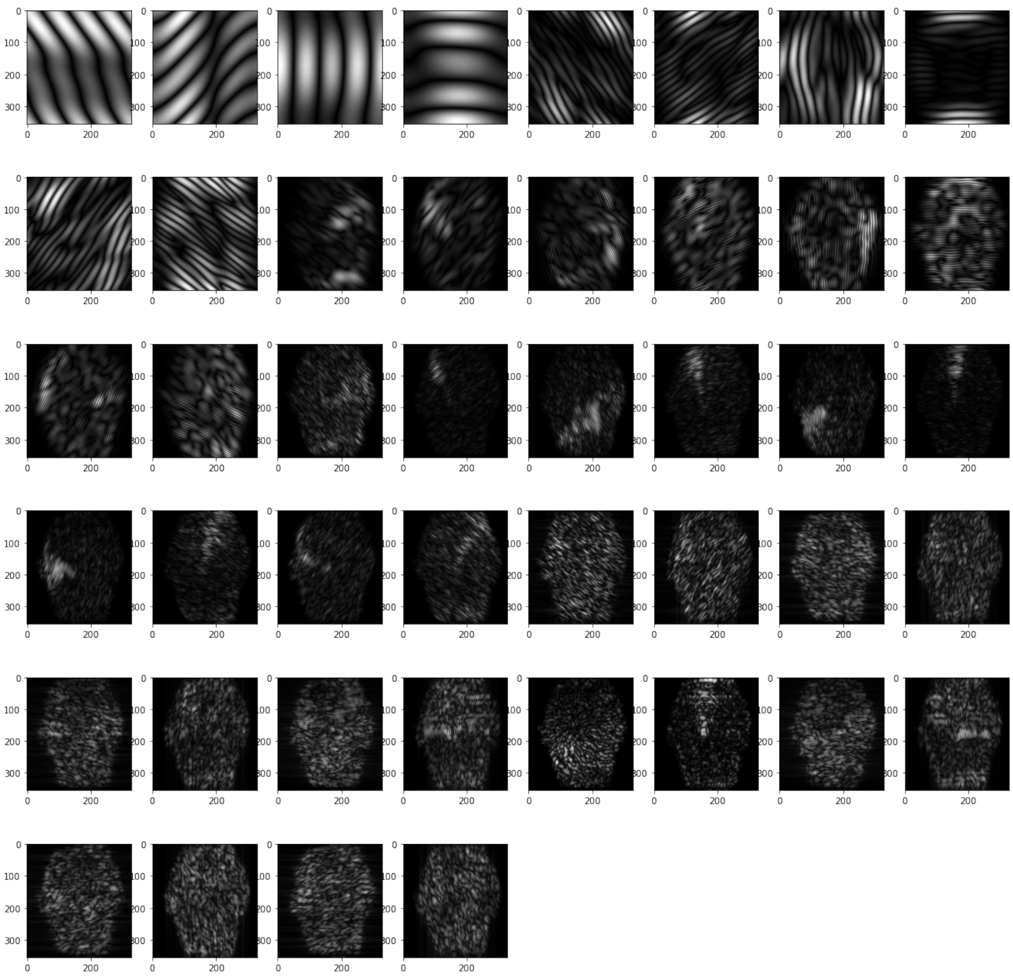
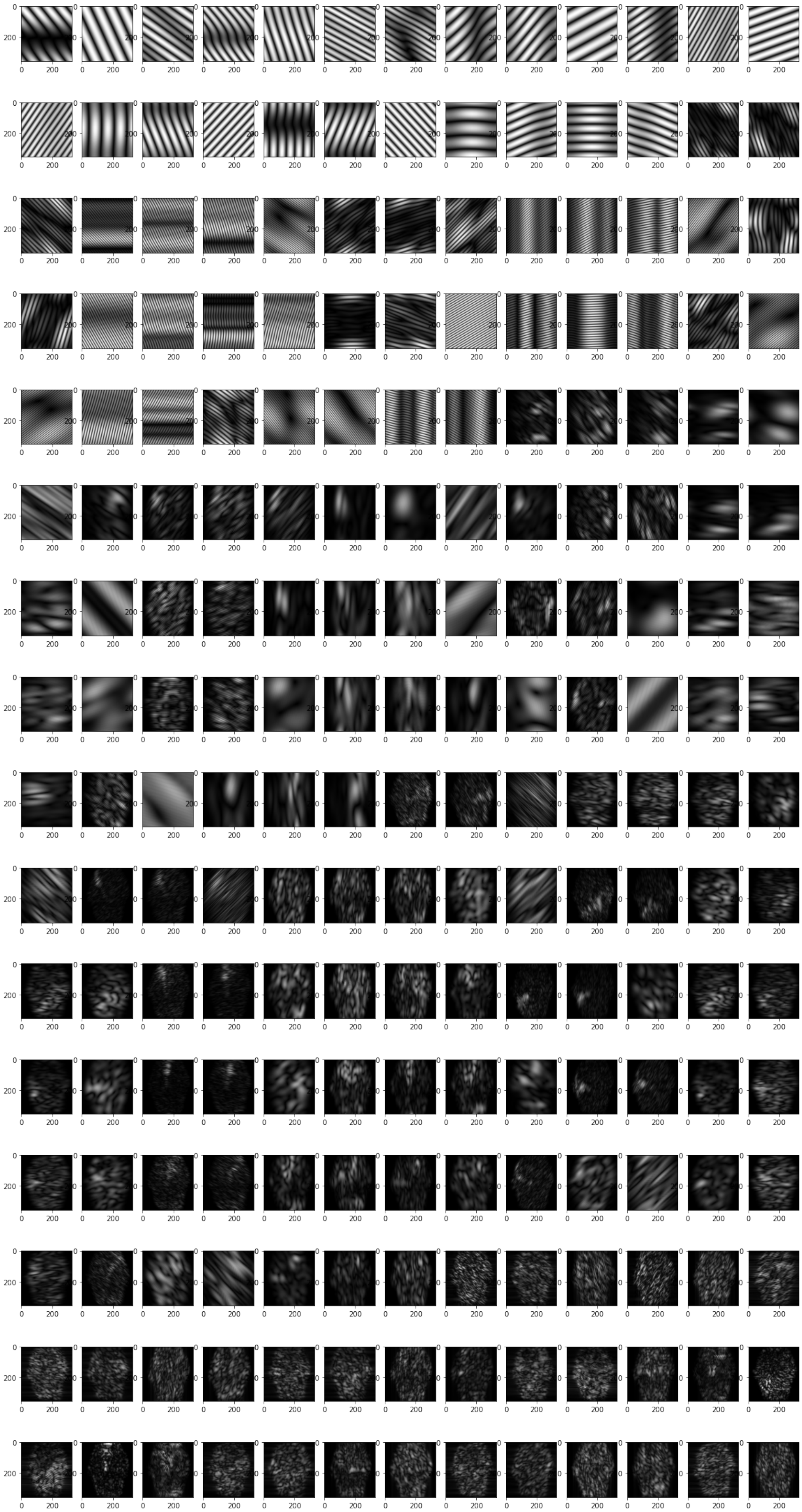


Figure 25: Level 2 Masks of new Meyer wavelet; note that many masks seem to be empty but aren't

Figure 26: Output of new Level2 masks on a fingerprint



9 Explainability

- We can see from the results that the number of trainable parameters for the scattering wavelet transform model increases very much on increasing the depth (J) more than 5 times which is a very huge number this is because each additional layer introduces more filters and coefficients, leading to a higher model complexity as explained in the paper by Stephane Mallat that depth increases the coefficients a lot.
- On increasing the angle index (L), we can observe that the trainable parameters increase but not more than 2 times this is because increasing the angle index in scattering wavelet transform introduces more orientations of the wavelet filters, and each orientation contributes additional learnable parameters. This increase in parameters allows the model to capture more complex and fine-grained features but also increase computational complexity and the potential risk of overfitting.
- If we increase the number of convolutional blocks it decreases the number of parameters because convolutional blocks decreases the size of feature maps thus reducing the number of weights required.
- We have trained the model on only 20 epochs because of fewer computational resources however if we would have trained on more number of epochs then the accuracy would have been increased to over 95 percent with the tradeoff between number of epochs and number of trainable parameters so that we don't do overfitting.
- In Shearlet system, we think that cosine similarity and Euclidean distance are not working because the features are very generic. They are in a direction but there is a difference between 2 images of the same fingerprint and the effective differences add up to make it similar to the other people's fingerprints as well.

10 Scalability

- The scattering wavelet model is a very basic model therefore it can be applied for a huge dataset. For a huge dataset we will need more number of trainable parameters such that the model does not overfit for the training data. to do this we can increase the depth as well as the angle index of scattering wavelet transform so that it can also capture more finer details of fingerprints.
- The idea of level 2 shearlet transform gives us many masks. Each of these masks define feature in a very specific direction and they can be used to identify a lot of features of the fingerprint. Using them for identification is possible. They give far too many features. I have worked on selection of specific features of these which do not have energy zero. However, I still have to see for further refinement. We can also try sending all these features as channels of an image into a CNN or Siamese network.

11 Planned Work

We plan to incorporate some Machine Learning Techniques to learn the features that are used to detect and verify fingerprints. Given below are 2 of the model techniques that we plan to implement :

- Siamese Model : This type of neural network is used to compare two inputs. It is particularly useful for fingerprint dataset models because it can be used to learn to extract features from fingerprints that are invariant to changes in rotation, translation, and scale. This makes it possible to build fingerprint matching systems that are more robust to noise and variations in the way that fingerprints are captured. In addition, Siamese networks can be used to learn to distinguish between genuine and fraudulent fingerprints. This is useful for building fingerprint-based authentication systems that are more secure.
- Few Shot Learning : Few-shot learning is particularly useful for fingerprint dataset models because fingerprint datasets can be difficult and expensive to collect and label for example our dataset has only 6 images per person, With few-shot learning, it is possible to train accurate fingerprint dataset models with only a small number of labeled examples in our case 6.

Other than these we will try to use scattering wavelet transform (SWT) which was explained in the lectures by Stephane Mallat. SWT works by decomposing the fingerprint image into a series of wavelet scattering coefficients. These coefficients capture the local texture and frequency information of the image. SWT is also able to capture long-range dependencies in the image, which is important for distinguishing between different fingerprint patterns.

12 Acknowledgement and Conclusion

Therefore the interim project report is competed. Thanks a lot to Professor Gadre to give such a great learning project.