International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2019 IJSRCSEIT | Volume 4 | Issue 8 | ISSN : 2456-3307
DOI : https://doi.org/10.32628/IJSRCSEIT

# Study of WCNN:NIR-VIS Face Recognition

**Swapnil Aknurwar[1], Shubham Thapa[2], Shubhakar Kumar Pandit[3], Nikita Nair[4], Prof. Jayashree Patil Chaudhari[5]**

[1,2,3,4] Student, Department of Computer Engineering, Dr. D. Y. Patil School of Engineering, Lohegoan, Savitribai Phule Pune University, Pune, Maharashtra, India

[5] Professor, Department of Computer Engineering, Dr. D. Y. Patil School of Engineering, Lohegoan, Savitribai Phule Pune University, Pune, Maharashtra, India

## ABSTRACT

In this survey paper we have been studying the WCNN Near infrared-visible (NIR-VIS) heterogeneous face recognition (HFR) refers to the process of matching NIR to VIS face images. The current heterogeneous methods try to extend VIS face recognition methods to the NIR spectrum by synthesizing VIS images from NIR images. It refers to matching a sample face image to a gallery of face images taken from alternate imaging modality. The major challenge of heterogeneous face recognition found in the great discrepancies between different image modalities. This survey paper having high resolution for heterogeneous face synthesis as complementary combination of two or more components. The painting component synthesizes and in paints VIS image textures from NIR image textures. The correction component maps any pose in NIR images to a frontal pose in VIS images, resulting in paired NIR and VIS textures. A warping procedure is developed to integrate the two components into an end-to-end deep network. A discriminator and wavelet based discriminator are being designed to supervise intra-class variance and visual quality respectively.

**Keywords :** Heterogeneous Face Recognition, Deep Neural Networks, VIS-NIR Face Matching, Feature Representation.

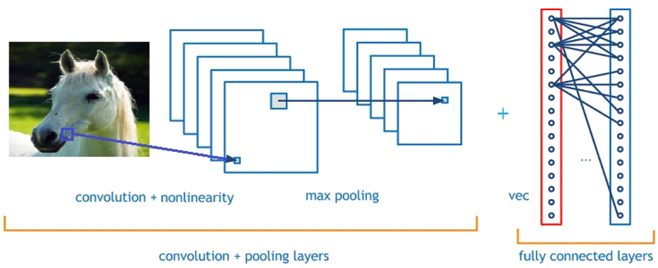
## I. INTRODUCTION

Face images can be captured by different acquisition systems like visible light cameras capture visible light (VIS) images while near infrared (NIR) images are captured by infrared imaging devices. The images can be captured at daytime or nighttime so illumination conditions also differ. Such types of images are known as heterogeneous images and heterogeneous face matching refers to matching face images across different modality. The heterogeneous (NIR-VIS) face images. In many face matching system, one of the difficult task is to match the heterogeneous face images. Many applications such as E passport, video surveillance, and photo based identification requires heterogeneous face matching because, in these applications probe face images and gallery face images are of different modalities such as NIR (Near Infrared), VIS (Visible Light), Matching faces across different sensing modalities raises the problem of heterogeneous face recognition (HFR) or cross-modality face recognition. By remarkable difference in sensing processes, heterogeneous images of the same subject have a large appearance variation, which has distinguished HFR from regular visual (VIS) face recognition.

In this paper, the two aforementioned problems are solved by a novel Wasserstein CNN (WCNN) architecture [1]. The WCNN employs a single network structure to map both NIR and VIS images to a compact Euclidean feature space such that the NIR and VIS images in the embedding space directly correspond to the face similarity. The WCNN[1] is composed of three key components in an end-to-end fashion. First of all, inspired by the observations and results indicating that the facial appearance is composed of identity information and variation information (e.g. lightings, poses, and expressions). We are trying to study and divide the high level layer of the WCNN into two orthogonal subspaces that contain modality-invariant identity information and modality-variant spectrum information.

## II. METHODOLOGY

CNNs is a Neural Networks that have proven very effective in areas such as image recognition and classification. CNNs are a type of feed-forward neural networks made up of many layers. CNNs consist of filters or kernels or neurons that have learnable weights or parameters and biases. Each and every filter takes some inputs, performs convolution and optionally follow it with a non-linearity . A CNN architecture can be seen as shown in Fig.1. The structure of CNN contains Convolutional, pooling, Rectified Linear Unit (ReLU), and Fully Connected layers.




**Convolutional Layer**: Convolutional layer performs the core building block of a Convolutional Network that does most of the computational heavy lifting. The primary purpose of Convolution layer is to extract features from the input data which is an image. Convolution preserves the spatial relationship between pixels by learning image features using small squares of input image. The input image is convoluted by employing a set of learnable neurons. This produces a feature map or activation map in the output image and after that the feature maps are fed as input data to the next convolutional layer.

**Pooling Layer:** Pooling layer reduces the dimensionality of each activation map but continues to have the most important information. The input images are divided into a set of non-overlapping rectangles. Each region is down-sampled by a non-linear operation such as average or maximum. This layer achieves better generalization, faster convergence, robust to translation and distortion and is usually placed between convolutional layers.

**ReLU Layer:** Basically ReLU is a non-linear operation and includes units employing the rectifier. It's an element wise operation that means it is applied per pixel and reconstitutes all negative values in the feature map by zero. In order to understand how the ReLU operates, we assume that there is a neuron input given as x and from that the rectifier is defined as f(x)= max (0, x) in the literature for neural networks.

**Fully Connected Layer:** Fully Connected Layer (FCL) term refers to that every filter in the previous layer is connected to every filter in the next layer. The output from the convolutional, pooling, and ReLU layers are embodiments of high-level features of the input image. The goal of employing the FCL is to employ these features for classifying the input image into various classes based on the training dataset.

## III. RELATED WORK

We are studying high level features of Deep Convolutional Neural Networks trained on visual spectra images are potentially domain independent and can be used to encode faces sensed in different image domains. A general framework for WCNN[1] Heterogeneous Face Recognition is proposed by adapting Deep Convolutional Neural Networks low level features in, so called, “Domain Specific Units”. This following adaptation using Domain Specific Units allow the learning of shallow feature detectors specific for each new image domain. Later on, it handles its transformation to a generic face space shared between all image domains. Heterogeneous Face Recognition (HFR) consists in matching faces from different image modalities.

## IV. EXSTING SYSTEM

We are studying investigated facial recognition algorithms to identify faces by extracting landmarks or features from an image. WCNN[1] algorithm may analyze the relative position, size, and/or shape of the eyes, nose, cheekbones and jaw. Those features are further used to search for other images with matching features.So this algorithms normalize a dataset of face images and then compress the face data, only saving the data in the image that is useful for face detection. A image is examined then compared with the face data. Previously successful systems are based on matching techniques applied to a set of salient facial features, providing a sort of compressed face representation. Hidden Markov model is a statistical model in which the system being modeled is assumed to be a Markov process with unobserved state. So as HMM can be considered as the simplest dynamic Bayesian network.

**DISADVANTAGES:**

- There are many extracted features leading to the variability of images of a single face that add to the complexity of the recognition problem if they cannot be avoided by careful design of the capture situation.
- A facial biometric security system can drastically improve your security because every individual who enters your premises will be accounted for. Any trespassers will be quickly captured by the recognition system and you would be alerted promptly.

## V. SYSTEM OVERVIEW

In this paper we have studied how Machine Learning Convolutional Neural Network algorithm were used to decide matching face recognition is to identify through input NIR and VIS imaging that identifying real or fake image with the help of convolutional neural network algorithm. We systemically evaluate the proposed WCNN[1] approach against traditional methods and deep learning methods on three recently published NIR-VIS face databases: the CASIA NIR-VIS 2.0 database, the Oulu-CASIA NIRVIS database and the BUAA-VIS NIR database.

Fig. 1 shows the proposed system architecture Wasserstein distance is used to measure the difference between NIR and VIS distributions in the modality invariant subspace (spanned by matrix W).

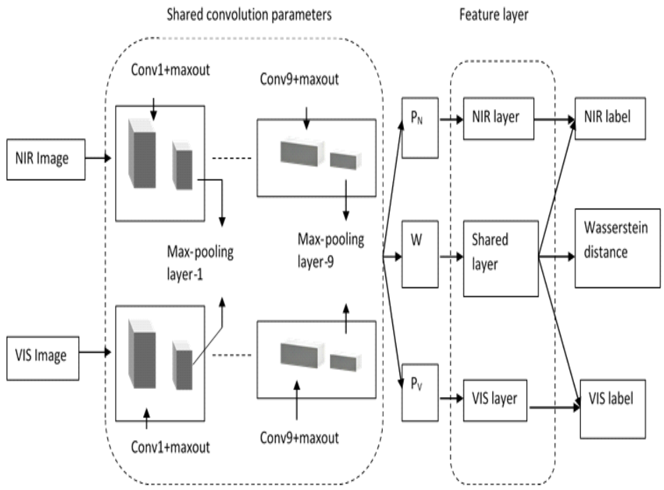



**Fig. 1.** Proposed System Architecture

## VI. LITERATURE SURVEY

| SR.NO | Paper name/author/yr | Functionality | Advantages | Disadvanta-ges |
|---|---|---|---|---|
| 1 | Wasserstein CNN: Learning Invariant Features for NIR-VIS | Wasserstein CNN approach that uses only one network to project both NIR and VIS | The gap between the sensing | Limited availability of training samples |

| | | | | |
|---|---|---|---|---|
| | Face Recognition, Ran He , Xiang Wu , Zhenan Sun , Tieniu Tan | images to a compact euclidean space. The WCNN naturally combines subspace learning and invariant feature extraction | patterns of the vis and nir modalities. Over-fitting on small-scale training sets. | of cross modality face image pairs. |
| **2** | Heterogeneous Face Recognition Using Domain Specific Units, Tiago de Freitas Pereira, Andre Anjos, and Sebastien Marcel | high level features of Deep Convolution Neural Networks trained on visual spectra images are potentially domain independent and can be used to encode faces sensed in different image domains. A general framework for Heterogeneous Face Recognition is proposed by adapting Deep Convolutional | It handles its transformation to a generic face space shared between all image domains. | will focus on the analysis on what such feature detectors are learning for each image domain. |
| **3** | Learning a High Fidelity Pose Invariant Model for High-resolution Face Formalization, Jie Cao, Yibo Hu, Hongwen Zhang, Ran He, Zhenan Sun | Face formalization refers to predicting the frontal view image from a given profile. It is an effective preprocessing method for pose-invariant face recognition. Formalized profile faces can be directly used by general face recognition methods without retraining the recognition models. | The prerequisite of warping is decomposed into dense correspondence field estimation and facial texture map recovering, | The prerequisite of warping is decomposed into dense correspondence field estimation and facial texture map recovering, |
| **4** | Cross-spectral Face Completion for NIR-VIS Heterogeneous Face Recognition , Ran He, Jie Cao, Lingxiao Song, Zhenan Sun, Tieniu Tan | Wasserstein convolutional neural network (WCNN) approach for learning invariant features between near-infrared (NIR) and visual (VIS) face images (i.e., NIR-VIS face recognition. | To avoid the over-fitting problem on small-scale heterogeneous face data, a correlation prior is introduced on the fully-connected WCNN layers to reduce the size of the parameter space. | Future work is whether this IDR framework can be useful for other heterogeneous or cross-modal problems, e.g., cross-sensor iris recognition and face recognition |

## VII. ADVANTAGES

- Because of weaker scattering and absorption, NIR light can penetrate deeper into. In excitation light in the visible (VIS) region cannot reach the imaging target in tissues in the conventional VIS-VIS imaging.
- The main advantage of Convolutional Neural Network compared to its predecessors is that it automatically detects the important features without any human supervision.
- The large-scale VIS dataset is helpful for VIS face recognition; it has limited benefits for HFR if only a small-scale NIR dataset is available.
- The accuracy requirement of face-based biometric recognition, by taking advantages of the recent NIR face technology while allowing the use of existing VIS face photos as gallery templates.

## VIII. CONCLUSION

Hence, we have studied how WCNN NIR-VIS FACE RECOGNITION across-spectral joint dictionary learning technique to reconstruct images between the NIR and VIS domain. The WCNN naturally combines subspace learning and invariant feature extraction into a CNN, and divides the high-level layer of the CNN into two orthogonal subspaces that contain modality-invariant identity information and modality-variant light spectrum information. The Wasserstein distance is used to measure the difference between heterogeneous feature distributions, and it is effective at reducing the sensing gap.

## IX. REFERENCES

[1]. Ran He , Senior Member, IEEE, Xiang Wu , Zhenan Sun , Member, IEEE, and Tieniu Tan,” Wasserstein CNN: Learning Invariant Features for NIR-VIS Face Recognition”

[2]. Tiago de Freitas Pereira “Heterogeneous Face Recognition Using Domain Specific Units” Journal of Latex Class Files, vol. 14, no. 8, August 2015.

[3]. Jie Cao “Learning a High Fidelity Pose Invariant Model for High-resolution Face Formalization” 32nd Conference on Neural Information Processing Systems (Nuri’s 2018), Montréal, Canada.

[4]. Benjamin S. Riggan “Synthesis of High-Quality Visible Faces from Polarimetric Thermal Faces using Generative Adversarial Networks” International Journal of Computer Vision manuscript.

[5]. Zhenan Sun “Cross-spectral Face Completion for NIR-VIS Heterogeneous Face Recognition” Journal of Latex Class Files, January 2017.

**Cite this article as :**

Swapnil Aknurwar, Shubham Thapa, Shubhakar Kumar Pandit, Nikita Nair, Prof. Jayashree Patil Chaudhari, "Study of WCNN:NIR-VIS Face Recognition", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 4 Issue 8, pp. 81-85, September-October 2019.
Journal URL : http://ijsrcseit.com/CSEIT194820